



# Externalities in a life cycle model with endogenous survival<sup>☆</sup>

Michael Kuhn<sup>a,\*</sup>, Stefan Wrzaczek<sup>b,a</sup>, Alexia Prskawetz<sup>c,a</sup>, Gustav Feichtinger<sup>b,a</sup>

<sup>a</sup> Austrian Academy of Science, Vienna Institute of Demography/Wittgenstein Centre, Wohllebengasse 12-14, A-1040 Vienna, Austria

<sup>b</sup> Institute of Mathematical Methods in Economics (Research Group on Operations Research and Control Systems), Vienna University of Technology, Argentinierstrasse 8, 1040 Vienna, Austria

<sup>c</sup> Institute for Mathematical Methods in Economics (Research Group on Economics), Vienna University of Technology, Argentinierstrasse 8, 1040 Vienna, Austria

## ARTICLE INFO

### Article history:

Received 22 November 2010

Received in revised form

31 May 2011

Accepted 11 August 2011

Available online 7 September 2011

### Keywords:

Demand for health

Externality

Life cycle-model

Optimal control

Tax-subsidy

Value of life

## ABSTRACT

We study socially vs individually optimal life cycle allocations of consumption and health, when individual health care curbs own mortality but also has a spillover effect on other persons' survival. Such spillovers arise, for instance, when health care activity at aggregate level triggers improvements in treatment through learning-by-doing (positive externality) or a deterioration in the quality of care through congestion (negative externality). We combine an age-structured optimal control model at population level with a conventional life cycle model to derive the social and private value of life. We then examine how individual incentives deviate from social incentives and how they can be aligned by way of a transfer scheme. The age-patterns of socially and individually optimal health expenditures and the transfer rate are derived. Numerical analysis illustrates the working of our model.

© 2011 Elsevier B.V. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/3.0/).

## 1. Introduction

Starting from the seminal work by Grossman (1972), economists have applied the life cycle model to examine how individuals allocate health care and consumption over their life course and what this implies for their health status, mortality and ultimately for longevity (e.g. Ehrlich and Chuma, 1990; Ehrlich, 2000). A related line of literature employs life cycle models to assess an individual's willingness to pay for survival – the value of life – and how it evolves over the life course (e.g. Shepard and Zeckhauser, 1984; Rosen, 1988; Johansson, 2002; Murphy and Topel, 2006). Both strands of the literature typically take a positive approach, i.e. they examine the determinants of individual health care choices and valuations but do not question their efficiency. In most of the

models, efficient life cycle choices are guaranteed anyway, as the individual (i) faces perfect markets, in particular, a perfect annuity market, and (ii) acts as an isolated decision-maker who is not linked to other individuals (contemporary or future).<sup>1</sup> In the real world, neither of the (implicit) assumptions in (i) and (ii) is likely to hold. Markets are typically imperfect or even missing so that externalities may arise. Likewise, individuals are linked to others not only through altruistic ties but also through externalities. Thus, efficiency is by no means guaranteed. In this paper, we seek to shed some light on the implication of intra- and intergenerational externalities on the efficiency of the life cycle choices over health care and consumption as well as on the underlying valuations. Specifically, we focus on the effects of spill-overs related to medical spending.

### 1.1. Externalities in the provision of health care

Typically, individual mortality not only depends on the individual's own consumption of health care but also on the level of aggregate 'activity' within a health care system. One could think of

<sup>☆</sup> We are grateful to an anonymous referee and the participants of various seminars and conferences for very helpful comments. All responsibility for remaining errors lies with us. This research was partly financed by the Austrian Science Fund under contract number I476-N13 (Endogenous heterogeneity and periodicity in dynamic optimisation) and by the European Commission under grant SSH-2007-3.1.01-217275 (Long-Run Economic Perspectives of Ageing Societies).

\* Corresponding author.

E-mail addresses: [michael.kuhn@oeaw.ac.at](mailto:michael.kuhn@oeaw.ac.at) (M. Kuhn), [wrzaczek@server.eos.tuwien.ac.at](mailto:wrzaczek@server.eos.tuwien.ac.at) (S. Wrzaczek), [afp@econ.tuwien.ac.at](mailto:afp@econ.tuwien.ac.at) (A. Prskawetz).

<sup>1</sup> Note, however, a number of important departures from these assumptions: (i) Shepard and Zeckhauser (1984), Ehrlich (2000) and Johansson (2002) examine the case in which individuals do not have access to an annuity market. Davies and Kuhn (1992) and Philipson and Becker (1998) examine moral hazard on annuity markets. (ii) Basu and Meltzer (2005) examine a model in which the individual cares for a contemporary family member (partner), whereas Birchenall and Soares (2009) and Kuhn et al. (2010) allow for altruism towards descendants.

a number of ways in which aggregate activity may either enhance or compromise individual efforts to reduce mortality. First, medical research has identified a positive relationship between volume of (surgical) activity and the quality of care, frequently measured by (lower) mortality (for an overview see Phillips and Luft, 1997). Thus, the effectiveness of individual health care increases in aggregate activity. Conversely, by contributing towards aggregate activity an individual also enhances the effectiveness of health care targeted at others. As long as these benefits are not internalised in the price of care – and there is no reason to believe they are, whatever the pricing arrangements – such spillovers constitute a positive externality.

Second, in as far as the provision of health care contains public good aspects, such as the provision of medical facilities, hospital bed capacity or emergency services, higher levels of aggregate health care spending may translate into a lower mortality risk at individual level. Similarly, higher aggregate spending levels may lead to greater scope for medical R&D or other quality enhancing activities that would not be lucrative in 'low spending' health care systems.<sup>2</sup>

Third, positive spillovers typically arise in the context of preventive activities. The most obvious example relates to vaccination, where individual mortality decreases in the degree to which the population is vaccinated against an infectious disease (for an overview see Philipson, 2000). The same applies to antimicrobial treatment of infectious disease, which is curative from the individual's perspective but also prevents further infections. Other preventive activities that lower both own and other people's mortality include the installation of safety devices such as Anti Lock Breaking in automobiles or fire detectors in tenement flats.

Finally, we may think of measures related to public health such as the cleaning of sewerage, proper disposal of household waste or the reduction of air pollution. Cutler and Miller (2005) show that in the early 20th century nearly half of the total mortality reductions in major US cities can be attributed to the introduction of clean-water technologies, i.e. the filtration and chlorination of water supplies.<sup>3</sup> 'Pure' public health measures constitute a polar case, where mortality reductions are exclusively due to cumulative expenditure.<sup>4</sup> Such a situation is equivalent to a public goods problem. But even in less extreme cases, the problem of private underprovision arises as long as a part of private health expenditure flows towards a public good (i.e. communal reductions in mortality).<sup>5</sup>

All of the above examples relate to positive spillovers, where higher activity translates into lower mortality. However, in a number of circumstances the converse may be true: aggregate activity may increase individual mortality. Negative spillovers could arise from congestion effects or from microbial resistance against antibiotics. Excessive demand for health care may lead

to congestion in the presence of capacity constraints. Hospital crowding, for instance, is likely to hike up mortality due to increased infection risks or due to over-stretched clinical staff lowering the attention afforded to the care of each individual patient and being more prone to committing medical errors.<sup>6,7</sup> Moreover, it is well known that microbes tend to develop resistance against antimicrobial treatments. The probability that a resistant microbial strain develops increases in the level of exposure. Thus, individual use of antibiotics tends to curb individual mortality but may, in aggregate, lead to an increased mortality risk due to microbial resistance.<sup>8</sup> In the case of negative spillovers, there is a tendency towards an excessive consumption of care.

The empirical relevance of all of the above mechanisms has been well documented (in the literature referenced). While we are unaware of empirical evidence as to the distortionary effects of these spillovers on the level and pattern of individual health expenditure, in the light of their prominence, we would expect these effects to be of a non-trivial magnitude.<sup>9</sup>

## 1.2. Life cycle implications of health-related externalities

The general implications of health-related externalities are straightforward enough: under- (over-)consumption of health care in the case of positive (negative) externalities. The life cycle aspects of the problem, however, are far more intricate. First, a distinct life cycle pattern of health care spending translates into a distinct pattern at which externalities are generated. Second, through its influence on mortality and ultimately on life expectancy the externality generates an important feedback. This is because changes in life expectancy have a bearing on the individual's aggregate wealth and on the need to spread this wealth over a life span of changing length. Third, the extent/value of the externality is endogenous as it depends on the size and age-structure of the population, the latter being determined by age-specific mortality. Finally, a transfer policy aimed at internalising the spill-overs needs to reflect the above properties and, therefore, gives rise to a particular age-schedule of the transfer.

We analyse these issues by combining two variants of a life cycle model with endogenous mortality, depending both on individual health expenditure and on a measure of aggregate health expenditure:

1. an age structured optimal control model, where a social planner maximises welfare at population level (i.e. individual utilities aggregated over time and age groups). This model determines the socially optimal pattern of consumption and health expenditure.

<sup>6</sup> Black and Pearson (2002) discuss the problems related with a recent bout in hospital congestion in the UK. Although congestion is predominantly a problem of public health care systems (see, however, footnote 7), this nevertheless implies the presence of negative spillovers within the system.

<sup>7</sup> Clement et al. (2008) use DEA techniques to identify congestion, i.e. the production of undesirable outputs (higher risk adjusted mortality for five conditions) together with desired outputs (treatments). They find for year 2000 data that 67% of US hospitals were experiencing some level of congestion amounting to an average efficiency loss of 13%.

<sup>8</sup> The positive correlation between intensity of antibiotic use and microbial resistance is well documented empirically (e.g. Cohen and Tartasky, 1997; Easterlin, 1999). See also Laxminarayan and Brown (2001) for an economic epidemiological model of optimal drug use in the presence of antibiotic resistance.

<sup>9</sup> A number of the externalities discussed (positive and negative) could be avoided through appropriate supply side policies (e.g. hospital congestion). For the purpose of this paper we take their absence as given and ask how the demand for health care could be adjusted to optimal levels.

<sup>2</sup> Murphy and Topel (2007), for instance, model an R&D race for a pharmaceutical innovation and show that the overall probability of innovation increases in the share of the social value that the winning firm is able to capture. In our model the prize for innovation would correspond to the winner's share of aggregate health expenditure, thus establishing a link between aggregate health expenditure and individual mortality.

<sup>3</sup> Watson (2006) studies the impact of public sanitation interventions in US Indian Reservations on the child mortality of Native Americans in the US as opposed to White infants. She finds that they were quite effective in reducing the mortality gap despite a sizeable externality on the health of neighbouring White children.

<sup>4</sup> Easterlin (1999) argues that, indeed, most of the historical reductions in mortality due to preventive measures, vaccination and antimicrobials are not attributable to the market for reason of various forms of externalities.

<sup>5</sup> An alternative but analogous interpretation is one in which health care is a good with (positive) network externalities.

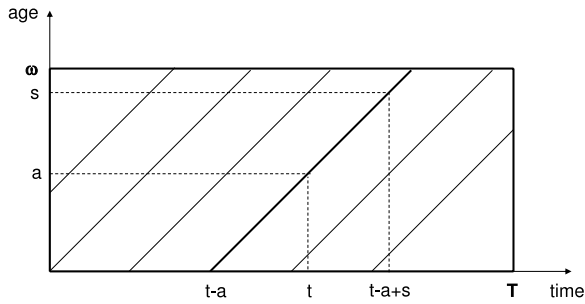


Fig. 1. Lexis diagram.

2. a 'conventional' life cycle model, where an individual maximises life time utility without altruistic concerns. This model determines the individual pattern of consumption and health expenditure.

Solving and simulating models (1) and (2) and comparing the respective patterns of consumption and health care we can deduce conclusions about the inefficiencies in individual behaviour and for which age-groups they arise due to the externality. We derive value of life expressions for the two models, a comparison of which allows to pinpoint the age-specific degree of inefficiency. Furthermore, we combine the two models in order to derive a transfer scheme that, if targeted at the individual, restores the first-best allocation.

To our knowledge, we are the first to integrate externalities into a life cycle model with endogenous survival.<sup>10</sup> As it will emerge, this is no trivial undertaking if the spillovers extend across different age-groups (or, indeed, cohorts) at any given point in time. By their very nature, the analysis and evaluation of such effects requires a model of the full population such as model (1), and therefore stretches beyond what could be achieved within an individual life cycle model alone. The latter allows to analyse the behaviour of a given cohort *along* the time path.<sup>11</sup> However, by construction individual life cycle models are not amenable to an analysis of cross-cohort effects. Our approach provides a consistent and tractable way of analysing such effects. Technically, the age-structured control model (1) differs from the life cycle model (2) in that the control variables – consumption and health care – and state variables – assets and population size/survival probability – vary in two dimensions – age and time – instead of a single dimension age = time.<sup>12</sup> The distinction can be illustrated with reference to the Lexis-diagram in Fig. 1, depicting the life experience of cohorts in time versus age. The 45° line represents cohort lines (e.g. a cohort born at  $t - a$  is of age  $a$  at time  $t$  and of age  $s$  at time  $t - a + s$ ).

Individual life cycle maximisation follows the cohort line disregarding other cohorts. The planner, in contrast, maximises over the whole Lexis-diagram. In the presence of spillovers, individual mortality develops along the cohort line not only

according to the individual's health spending  $h(a, t)$  alone but also according to the health spending  $h(\hat{a}, t)$ ,  $\hat{a} \neq a$  realised by other cohorts at each point in time,  $t$ . The first best result cannot be attained by individuals without correcting policy-measures.

Our main results are as follows. Unsurprisingly, the presence of positive (negative) spillovers implies that individuals spend too little (much) on health care and consequently consume too much (little). The divergence between individual and social incentives is rendered evident in the social as compared to the private value of health care. Individuals value the purchase of additional health care according to their private value of life (i.e. the discounted value of their expected remaining life time). In contrast, the planner values additional health care by the social value of life plus the value of the externality (i.e. the total value of mortality changes for the rest of the society induced by the individual's current health expenditure). The social value of an individual's life diverges from the private value of life because it incorporates the value of spillovers generated by the health care the individual consumes over its remaining life time. This constitutes a generalisation of the value of life to incorporate externalities.<sup>13</sup> We also consider the age-schedule of health expenditure depending on the relevant, private or social, value.

We derive the optimal transfer that fully internalises the spillover and achieves an efficient allocation of consumption and health care over the life cycle. Expressed as an ad-valorem subsidy/tax on health expenditure (with the resulting budget balance being reimbursed in a lump-sum fashion), the optimal transfer rate amounts to the proportion of the external value of health spending (current and future induced through the individual's survival) relative to its total value (including in addition the private value of life). We also show how the transfer varies with age.

Finally, a numerical analysis, based on US data akin to that used by Hall and Jones (2007), reveals an interesting asymmetry between positive and negative spillovers. If positive spillovers remain uncompensated, the inefficient under-spending leads to a significant reduction in life expectancy. Individual under-spending implies higher mortality as both too little is spent on individual survival (a direct effect) and too low a value of spillovers is generated (an indirect effect). In contrast, the inefficiency in the presence of negative spillovers leads to far less variation in life expectancy. On the one hand, over-spending leads to a direct reduction in individual mortality; on the other hand, it boosts mortality due to the negative spillovers. The result is a tread-mill effect, where individual efforts to curb mortality are neutralised, leading predominantly to a loss of consumption.

The remainder of the paper is organised as follows. Section 2 lays out the model. In Section 3, we derive the social optimum and characterise the social value of life. Individual choice is analysed in Section 4 where we present the individual optimum and then establish the transfer scheme that internalises the spillovers and induces the first-best allocation. In Section 5 we present numerical simulations to support the analytical results. The final section offers a discussion and outlook for further research.

## 2. The model

The dynamics of the population is described by the McKendrick equation (see Keyfitz and Keyfitz, 1997)

$$\begin{aligned} N_a + N_t &= -\mu(a, h(a, t), \bar{H}(t))N(a, t) \\ N(0, t) &= B(t), N(a, 0) = N_0(a). \end{aligned} \quad (1)$$

<sup>10</sup> Laxminarayan and Brown (2001) and Philipson and Mechoulam (2003) consider consumption externalities related to (pharmaceutical) products. While both studies examine the dynamic impact of internalising policies they do not consider a life cycle framework. Bolin et al. (2002) consider a life cycle setting where both an employee and an employer invest into the employee's health, which in turn affects her productivity. The resulting Nash-equilibrium is plagued by free-riding. Their set-up differs from ours as (i) life expectancy is fixed and exogenous; (ii) population structure plays no role; (iii) they do not derive an optimal policy.

<sup>11</sup> It is thus fit e.g. to analyse the life cycle effects of missing annuity markets as in Shepard and Zeckhauser (1984) or Ehrlich (2000).

<sup>12</sup> Hall and Jones (2007) consider a discrete time formulation of an age-structured life cycle model in order to simulate the development of health care-spending in the US. However, they do not consider externalities. Neither do they compare individual and social choices.

<sup>13</sup> Birchenall and Soares (2009) and Kuhn et al. (2010) provide another generalisation, where the value of progeny is incorporated into the value of life.

The state variable  $N(a, t)$  represents the number of  $a$ -year old individuals at time  $t$ . The left hand side denotes the change of  $N(a, t)$  in an infinitesimal small step of time  $\Delta$  along the 45° line in the Lexis diagram in Fig. 1, i.e. from  $(a, t)$  to  $(a + \Delta, t + \Delta)$ . The subindices  $a$  and  $t$  correspond to partial derivatives, i.e.  $N_a$  and  $N_t$  denote the partial derivatives of  $N(a, t)$  with respect to  $a$  and  $t$  respectively.  $N(a, t)$  decreases by the age-specific mortality rate  $\mu(a, h(a, t), \bar{H}(t))$ , which depends on age  $a$  and can be reduced instantaneously by providing to the individual an age-specific amount  $h(a, t)$  of health care (or other goods serving to curb mortality). Here,  $h(a, t)$  is an age-structured control variable in our model. Individual mortality also depends on health care spending at population level,  $\bar{H}(t)$ .  $N_0(a)$  describes the initial age distribution of the population and  $B(t)$  equals the exogenous number of newborns at time  $t$ , which allows for a constant, a growing or a shrinking population.

We assume that the mortality rate  $\mu(a, h(a, t), \bar{H}(t))$  is a strictly concave function in  $h$  and  $\bar{H}$ , satisfying an Inada condition with regard to  $h$ . Specifically, we assume

$$\mu(a, \cdot, 0) \leq \tilde{\mu}(a) \quad (\forall a); \quad \mu(a, 0, 0) = \tilde{\mu}(a) \quad (\forall a) \\ \mu_h(\cdot) < 0, \mu_{hh}(\cdot) > 0; \quad \mu_h(a, 0, 0) = -\infty \quad (\forall a),$$

where  $\tilde{\mu}(a)$  is the “natural” mortality rate resulting without any health care. Mortality cannot be worsened by investing in own health. We assume that the mortality rate depends on the aggregate level of health care, as measured by per capita consumption of health care  $\bar{H}(t)$ , (averaged across the full population) at time  $t$ .<sup>14</sup> The impact of aggregate activity on mortality, i.e.  $\mu_{\bar{H}}(a, h, \bar{H}) := \frac{\partial \mu(a, h(a, t), \bar{H}(t))}{\partial \bar{H}(t)}$ , is negative in the case of a positive externality (greater life cycle utility through lower mortality) and positive in the case of a negative externality (lower life cycle utility through higher mortality).

We can write per capita health expenditure  $\bar{H}(t) := \frac{H(t)}{\bar{N}(t)}$ , with

$$H(t) = \int_0^\omega h(a, t) N(a, t) da \\ \bar{N}(t) = \int_0^\omega N(a, t) da \quad (2)$$

denoting aggregate consumption of health care and total population at time  $t$  respectively. Here,  $\omega$  is the maximal age an individual can reach.<sup>15</sup> Aggregate health expenditure for age-group  $a$  at time  $t$  is given by the sum of individual expenditure,  $h(a, t) N(a, t) = \int_0^{N(a, t)} h_i(a, t) di$ . The impact of individual spending on average health expenditure is then given by  $\frac{\partial \bar{H}(t)}{\partial h_i(a, t)} = \frac{1}{\bar{N}(t)}$  which tends to zero for large populations. Thus, individuals rationally anticipate that they are unable to influence aggregate spending and an externality arises.

The second control variable is consumption  $c(a, t)$ . Consumption increases instantaneous utility according to the non-decreasing and concave function  $u(c(a, t))$ , which is assumed to satisfy the Inada condition  $\lim_{c \rightarrow 0+} u_c(c) = +\infty$ . The objective of the social planner is to maximise social welfare, defined as the sum of the instantaneous utilities of all individuals (total utilitarianism)

$$\int_0^T \int_0^\omega e^{-\rho t} u(c(a, t)) N(a, t) da dt.$$

The parameter  $\rho$  denotes the rate of time preference. Note that we can allow for an infinite planning horizon  $T = +\infty$  at this stage.

Finally we assume a budget constraint that is balanced for each cohort. This is expressed by the introduction of total wealth  $A(a, t)$  held by age-group  $a$  at time  $t$  and the following dynamics<sup>16</sup>:

$$A_a + A_t = rA(a, t) + (y(a) - c(a, t) - h(a, t))N(a, t) \\ A(0, t) = A(\omega, t) = 0 \quad \forall t \\ A(a, 0) = A_0(a), A(a, T) = A_T(a) \quad \forall a. \quad (3)$$

Here,  $r$  denotes the interest rate, assumed to be exogenous to the economy, and  $y(a)$  denotes an exogenous income accruing to an  $a$ -year old individual. We normalise the price of health care to one so that  $h(a, t)$  denotes both individual consumption of care and health expenditure. Each cohort is assumed to hold zero assets at the time of birth and death. By guaranteeing that each cohort spends precisely its own life time income, the cohort specific budget constraint allows us to compare steady-state allocations derived for the social planner with allocations derived for the individual (see Section 4). We thus rule out differences in total spending across cohorts which are due to discrepancies between the rate of time preference and the interest rate. Consider a situation where budgets are pooled across all cohorts. If  $\rho > r$ , for instance, the social planner would then not only (i) shift consumption to the beginning of the planning horizon for each individual cohort, but also (ii) shift consumption from future cohorts to present cohorts. Intuitively, we would like to account for shifts in consumption within cohorts (i) but not across cohorts (ii). Such an allocation rule also bears some intuitive appeal on equity grounds. Moreover, in the current paper we focus on the externality in survival. The effects related to such externalities can be isolated more clearly when there are as few as possible confounding effects.

The problem for the social planner is then to choose the age-specific schedule of consumption and health care to maximise social welfare. Formally, this is represented by the following dynamic age-structured optimisation problem with state variables  $A(a, t)$  and  $N(a, t)$  and control variables  $c(a, t)$  and  $h(a, t)$ ,<sup>17</sup>

$$\max_{c, h} \int_0^T \int_0^\omega e^{-\rho t} u(c(a, t)) N(a, t) da dt \\ \text{s.t. } N_a + N_t = -\mu(a, h(a, t), \bar{H}(t)) N(a, t) \\ N(0, t) = B(t), \quad N(a, 0) = N_0(a) \\ A_a + A_t = rA(a, t) + (y(a) - c(a, t) - h(a, t))N(a, t) \\ A(0, t) = A(\omega, t) = 0 \quad \forall t \\ A(a, 0) = A_0(a), A(a, T) = A_T(a) \quad \forall a \\ H(t) = \int_0^\omega h(a, t) N(a, t) da \\ \bar{N}(t) = \int_0^\omega N(a, t) da \\ \text{with} \\ \bar{H}(t) = \frac{H(t)}{\bar{N}(t)}.$$

Both parameters, time  $t$  and age  $a$ , are finite in our model since in general no transversality conditions are available for

<sup>14</sup> We use per capita consumption of health care in order to exclude scale effects as drivers of our results. We should stress, however, that the use of total consumption of health care as a measure of aggregate activity would not yield qualitatively different results.

<sup>15</sup> If chosen large enough the fixed upper limit to life time  $\omega$  is no restriction to the model. In some approaches  $\omega$  is variable or a decision variable (e.g. Boucek et al., 2004; Ehrlich and Chuma, 1990). In our model endogenous life expectancy (below  $\omega$ ) emerges through the age-dependent mortality schedule  $\mu(a, h, \bar{H})$ .

<sup>16</sup> Analogously to the population dynamics,  $A_a$  and  $A_t$  correspond to partial derivatives of  $A(a, t)$  with respect to  $a$  and  $t$ . Thus (3) describes the change of  $A(a, t)$  in an infinitesimal small time step along the 45° line in the Lexis diagram.

<sup>17</sup> A similar modelling approach is pursued by Feichtinger et al. (2006) on capital replacement by firms, by Prskawetz and Veliov (2007) on intertemporal labour demand and training, and by Saglam and Veliov (2008) on maintenance and endogenous depreciation.



age-structured optimal control models for infinite parameters. However, as discussed earlier this is no restriction to the model. In the following section we derive and interpret the necessary optimality conditions of the above age-specific control problem.

### 3. The social optimum

To obtain necessary optimality conditions we apply the maximum principle for age-structured control models (Feichtinger et al., 2003). We define the Hamiltonian of the social welfare problem as follows<sup>18</sup>:

$$\mathcal{H} = u(c)N - \xi^N \mu(a, h, \bar{H})N + \xi^A(rA + (y - c - h)N) + \eta^H hN + \eta^N N \quad (4)$$

where we denote the adjoint variables that correspond to the state variables as follows:  $\xi^N(a, t)$  refers to age-group  $N(a, t)$ ,  $\xi^A(a, t)$  refers to assets  $A(a, t)$ ,  $\eta^H(t)$  refers to total health expenditure  $H(t)$ , and  $\eta^N(t)$  refers to total population  $\tilde{N}(t)$ . A full derivation of the social optimum can be found in the proof of Proposition 1 in Appendix A. Here, we restrain ourselves to reporting the necessary first order conditions with respect to consumption and health expenditure:

$$\mathcal{H}_c = u_c(c)N - \xi^A N = 0 \quad (5)$$

$$\mathcal{H}_h = -\xi^N \mu_h(a, h, \bar{H})N - \xi^A N + \eta^H N = 0. \quad (6)$$

Combining them we obtain

$$u_c(c) = -\xi^N \mu_h(a, h, \bar{H}) + \eta^H. \quad (7)$$

Condition (5) governs the distribution of consumption over the life cycle.<sup>19</sup> Condition (7) governs the choice of health expenditure. The LHS gives the foregone welfare (per capita) in marginal utility terms if health spending is increased by one unit for each member of age-group  $a$  at time  $t$ . As usual,  $\xi^N$  can be interpreted as a shadow price, indicating the increase of the value function (i.e. social welfare) for a small (marginal) increase of  $N(a, t)$ . In other words,  $\xi^N$  gives the social value in utility terms of an individual, implying a marginal social benefit of  $-\xi^N \mu_h$  due to life saving. The second term on the RHS in (7),  $\eta^H$ , represents the value of the externality generated by a marginal increase in health expenditure, where  $\eta^H > (<)0$  for positive (negative) externalities.

We can calculate the value of life as the willingness to pay for a small reduction of the mortality rate for age  $a$  at time  $t$ . To our knowledge this concept was first developed in a formal manner by Shepard and Zeckhauser (1984) (see also Rosen, 1988; Johansson, 2002) who apply it to an isolated individual decision-maker. In the presence of cross-cohort spillovers, a distinction needs to be made between the private and social value of life (PVOL and SVOL respectively). Denoting by  $V$  the value function corresponding to the planner's problem, we define the social value of an individual life as

$$\psi^S(a, t) := \frac{\partial V / \partial N}{\partial V / \partial A} = \frac{\xi^N(a, t)}{\xi^A(a, t)} = \frac{\xi^N(a, t)}{u_c(c(a, t))}. \quad (8)$$

We can then express the optimality condition in value terms as follows (a proof is given in Appendix A).

**Proposition 1.** *Socially optimal health care satisfies*

$$\frac{-1}{\mu_h(a, h, \bar{H})} = \psi^S(a, t) + \Theta(a, t) \quad (9)$$

$$\Theta(a, t) := \int_0^\omega \frac{u_c(c(\hat{a}, t))}{u_c(c(a, t))} \psi^S(\hat{a}, t) \frac{N(\hat{a}, t)}{\tilde{N}(t)} \frac{\mu_{\bar{H}}(\hat{a}, h, \bar{H})}{\mu_h(a, h, \bar{H})} d\hat{a}, \quad (10)$$

i.e. it equates the effective marginal cost,  $\frac{-1}{\mu_h(a, h, \bar{H})}$ , of saving the life of an individual aged  $a$  at time  $t$  to the SVOL for this individual plus the value of the externality,  $\Theta(a, t)$ , generated by the provision of care to an individual at age  $a$  and time  $t$ .

Before characterising further the SVOL let us turn to the value of the externality  $\Theta(a, t)$ , which is determined by the following factors: (i) the weighted sum over age-groups of the SVOL's for  $(\hat{a}, t)$  individuals, where the population share  $\frac{N(\hat{a}, t)}{\tilde{N}(t)}$  is used as weight; (ii) the relative effectiveness of aggregate spending in reducing the mortality of  $(\hat{a}, t)$  individuals as given by  $\frac{\mu_{\bar{H}}(\hat{a}, h, \bar{H})}{\mu_h(a, h, \bar{H})}$ ; and (iii) a conversion factor  $\frac{u_c(c(\hat{a}, t))}{u_c(c(a, t))}$ .<sup>20</sup> Obviously,  $\Theta(a, t) > (<)0$  if and only if  $\mu_{\bar{H}}(\cdot) < (>)0$ , implying that positive (negative) externalities translate into a positive (negative) value. The absolute value of an externality is then large if the spillovers are particularly effective (high  $|\frac{\mu_{\bar{H}}(\hat{a}, h, \bar{H})}{\mu_h(a, h, \bar{H})}|$ ) for large age groups (high  $\frac{N(\hat{a}, t)}{\tilde{N}(t)}$ ) composed of members with a high SVOL.

The following lemma provides a further characterisation of SVOL (a proof is given in Appendix B).

**Lemma 1.** *The SVOL*

$$\psi^S(a, t) = \psi^P(a, t) + \Omega(a, t)$$

is composed of the PVOL

$$\psi^P(a, t) = \int_a^\omega v(s, t - a + s) e^{-\int_a^s [r + \mu(s', h, \bar{H})] ds'} ds \quad (11)$$

$$v(a, t) := \frac{u(c(s, t - a + s))}{u_c(c(s, t - a + s)) + (y(s) - c(s, t - a + s) - h(s, t - a + s))} \quad (12)$$

and the discounted social valuation of net contributions towards the externality over the remaining life course

$$\Omega(a, t) := \int_a^\omega e^{-\int_a^s [r + \mu(s', h, \bar{H})] ds'} \times \left[ \frac{\Theta(t - a + s)}{\psi^S(t - a + s) + \Theta(t - a + s)} \right] ds. \quad (13)$$

The planner's valuation of an individual life thus comprises both the PVOL, counting directly towards social welfare, and the value she attaches to the individual's future contributions to the externalities. The PVOL consists of the (discounted) net private values of a life year as defined in (12) summed over the individual's remaining life time. The net private value of a life year consists of (i) gross consumer surplus  $\frac{u(c(\cdot))}{u_c(c(\cdot))}$  and (ii) the individual's net savings  $y(\cdot) - c(\cdot) - h(\cdot)$ . The PVOL corresponds precisely to the measure derived in individual life cycle models (e.g. Shepard and Zeckhauser, 1984; Rosen, 1988; or Johansson, 2002). In addition, the planner accounts for the individual's future

<sup>18</sup> From now on we omit  $a$  and  $t$  if they are not of particular importance.

<sup>19</sup> As is readily checked the optimal distribution of consumption implies  $\frac{u_c(c(a', t))}{u_c(c(a', t - a + a'))} = \frac{\xi^A(a', t)}{\xi^A(a', t - a + a')} = e^{-(\rho - r)(a' - a)}$ , the Euler condition for optimal saving.

<sup>20</sup> The conversion factor equals one if consumption is equalised across age-groups.

contribution towards social welfare through the spillovers from individual consumption of health care. The discounted value of these contributions  $\Omega(a, t)$ , again summed over the individual's remaining life time, is thus the second component of the SVOL.

The value of the individual's remaining contributions to the externality in (13) increases with the discounted sum over the remaining life time of the individual's net spending surplus over the average,  $h(s, t - a + s) - \bar{H}(t - a + s)$ , where the relative value of the externality

$$\frac{\Theta}{\psi^S + \Theta}, \quad (14)$$

serves as a weight. Note that the sign of the relative value depends on the nature of the externality and increases in absolute value with the strength of the externality  $\frac{\mu_{\bar{H}}(\bar{a}, h, \bar{H})}{\mu_h(a, h, \bar{H})}$ . Thus, the future contributions imply a social premium,  $\Omega(a, t) > 0$ , if the individual is spending in excess (short) of the average in the case of positive (negative) externalities. Conversely,  $\Omega(a, t) < 0$  if the individual is spending in excess (short) of the average and externalities are negative (positive).

Finally, substituting into (9) we can describe the socially optimal choice of health care by the first-order condition

$$\frac{-1}{\mu_h(a, h, \bar{H})} = \psi^P(a, t) + \Omega(a, t) + \Theta(a, t). \quad (15)$$

The planner's full valuation of health care (set equal to the effective cost of saving a life  $\frac{-1}{\mu_h(\cdot)}$ ) amounts to the sum of the private value of life, the value of the individual's future contributions towards the externality and the current value of the externality, as generated by current health care spending. In the following section, we will use this expression to compare the social incentive with the private incentive. Before that, we examine the dynamics of the system.

The optimal paths of consumption and health expenditure can be described as follows. The change in the consumption of a cohort born at  $t - a$  is given by<sup>21</sup>

$$c_a + c_t = \frac{u_c(c)}{u_{cc}(c)}(\rho - r) \quad (16)$$

and reflects the usual life cycle distribution of consumption. If the discount rate,  $\rho$ , equals the interest rate,  $r$ , the right hand side (RHS) is zero implying consumption smoothing over the whole life for each cohort. If  $\rho > r$  consumption decreases over the life cycle, reflecting the impatience of individuals. The opposite applies for  $r > \rho$ .

In Appendix C we derive the change in health expenditure over time and age as

$$\begin{aligned} h_a + h_t = & \underbrace{-\frac{\mu_{ha} + \mu_{h\bar{H}}\bar{H}_t}{\mu_{hh}}}_{(i)} \\ & - \underbrace{\frac{\mu_h}{\mu_{hh}} \left( (r + \mu) - \frac{1}{\psi^S} \left( v + \frac{\eta^H}{u_c(\cdot)}(h - \bar{H}) \right) \right)}_{(ii)} \\ & + \underbrace{\frac{1}{\mu_{hh}\psi^S} \frac{\eta^H}{u_c(\cdot)} \left( \frac{\eta_t^H}{\eta^H} - (\rho - r) \right)}_{(iii)}. \end{aligned} \quad (17)$$

The change in health expenditure can be decomposed into three effects. (i) describes the change in the marginal effectiveness of

health expenditure due to the individual's progressing age  $\mu_{ha}$  and due to changes over time in average spending  $\mu_{h\bar{H}}\bar{H}_t$ . Recalling that  $\mu_{hh} > 0$  (implying decreasing returns to health care spending at any given age) it follows that health expenditure tends to increase directly with age as long as the marginal effectiveness of health spending increases with age  $\mu_{ha} < 0$ . Typically, this tends to be true for ages up to the 60–70s with health expenditure having little impact on mortality before the onset of life threatening conditions. For the highest ages health expenditure is likely to become less effective in combatting mortality so that  $\mu_{ha} \geq 0$ . The impact of an increase in aggregate expenditure depends on the nature of the externality. Obviously, if aggregate expenditure raises (lowers) the effectiveness of individual spending, this should lead to an increase (depression) of individual expenditure.<sup>22</sup>

(ii) describes the impact of a change in the SVOL. Indeed, as we show in Appendix C the expression equals the rate of change  $\frac{\psi_a^S + \psi_t^S}{\psi^S}$ . On the one hand, SVOL increases at the effective interest rate  $r + \mu$ ; on the other hand, it falls as with the passing of a life year the social value of this year,  $v + \frac{\eta^H}{u_c}(h - \bar{H})$ , is lost. Note that the social value of each life year is made up of the private value  $v$  as defined in (12) and the social value of the individual's current net contribution to the externality. The latter amounts to the excess spending  $h - \bar{H}$  weighted with the relative value of the externality  $\frac{\eta^H}{u_c} = \frac{\Theta}{\psi^S + \Theta}$  as defined in (14).<sup>23</sup> Noting that  $-\frac{\mu_h}{\mu_{hh}} > 0$ , it follows that an overall increase (decrease) in SVOL tends to raise (depress) individual health expenditure over time.

(iii) describes the impact on health expenditure of a change over time in the relative value of the externality  $\left(\frac{\eta^H}{u_c}\right)_a + \left(\frac{\eta^H}{u_c}\right)_t$ . The relative value of the externality increases with the rate of change of the shadow price  $\eta^H$  and changes with consumption over time and age. Here, an increase in consumption with age, as for  $\rho < r$ , implies a reduction in the marginal utility of consumption  $u_c$ , thus rendering the externality relatively more valuable (as opposed to consumption). Consider, for instance,  $\frac{1}{\mu_{hh}\psi^S} \frac{\eta^H}{u_c(\cdot)} > 0$  as is typical for positive externalities. Then, health expenditure tends to increase with age if the value of the externality increases over time or if consumption increases with age, as for  $\rho < r$ .

#### 4. Individual choice and optimal transfer policy

In this section we derive the individual life cycle choices of health care and consumption and compare them with the social optimum. In contrast to the planner's problem, age and time are identical from the individual's perspective. Defining  $t_0 := t - a$  as the year of birth of an individual, we use the notation  $c(a) := c(a, t_0 + a) = c(a, t)$  for age-specific consumption, and analogously for all other age- and time-dependent variables  $h, A, \mu$  relating to an individual of age  $a$  at time  $t$ . In order to account for the fact that the individual optimises over its life course (corresponding to the 45° line in the Lexis diagram in Fig. 1) rather than the whole age-time rectangular (as the planner does) we use  $\hat{c}(a)$  for the individual optimisation instead of  $c_a + c_t$ . The time dependent level of aggregate health expenditure per capita is now indexed by  $t_0 + a$ . All variables and underlying functions are as defined before. For the moment consider a set-up in which individual life cycle choices are not subject to a correcting policy.

<sup>21</sup> The expression is obtained by calculating the directional derivative of (5) and using (29) from Appendix A.

<sup>22</sup> Examples are learning-by-doing, where the effectiveness of individual surgery increases with aggregate activity; and microbial resistance, where the effectiveness of individual pharmaceutical treatment is reduced. Indeed, one would then expect individual expenditure to be raised in the former case and to be lowered in the latter.

<sup>23</sup> The identity is readily verified from the second equality in (32) in Appendix A.

We follow Yaari (1965) in considering a set-up in which individuals can fully annuitise their wealth by trading actuarial notes at some return  $r + \mu(a, h(a), \bar{H}(t_0 + a))$ . We assume here that the individual fully anticipates the impact of changes in health care on the annuity return. While this assumption is clearly stylised, it allows us to abstract from longevity-related moral hazard (Davies and Kuhn, 1992; Philipson and Becker, 1998) as a confounding problem. Hence, individual wealth develops according to

$$\dot{A}(a) = (r + \mu(a, h(a), \bar{H}(t_0 + a)))A(a) + y(a) - c(a) - h(a) \quad A(0) = 0. \quad (18)$$

Assuming the absence of planned-for bequests, we obtain  $A(\omega) = 0$ . The probability of surviving to age  $a$  (modelled analogously to the social planner problem) equals

$$M(a) := \exp\left(-\int_0^a \mu(s, h(s), \bar{H}(t_0 + s)) ds\right).$$

We can thus express the life cycle problem for an individual born at  $t_0$  as follows

$$\begin{aligned} \max_{c, h} \int_0^\omega e^{-\rho a} u(c(a)) M(a) da \\ \text{s.t. } \dot{M}(a) = -\mu(a, h(a), \bar{H}(t_0 + a))M(a) \\ \dot{A}(a) = (r + \mu(a, h(a), \bar{H}(t_0 + a)))A(a) + y(a) - c(a) - h(a) \\ M(0) = 1, A(0) = 0, A(\omega) = 0. \end{aligned}$$

We have shown previously that individuals do not expect to affect health care spending per capita (as averaged across each age-group) and therefore take  $\bar{H}(t_0 + a)$  as given at each point in time. We show in Appendix D that the individual then chooses health expenditure such that

$$\frac{-1}{\mu_h(a, h, \bar{H})} = \psi^P(a). \quad (19)$$

The individual equalises effective marginal expenditure with the PVOL,  $\psi^P(a)$ , as defined in (11). The externality has a bearing on the effectiveness of individual health care spending and is therefore prone to shift individual spending patterns. However, individuals do not take into account the benefit (harm) they bestow on others in the case of positive (negative) spillovers. A comparison with (15) shows immediately that individuals fail to take account of both the value of the individual's future contributions towards the externality and the current value of the externality. Thus, they will under-spend (over-spend) in the case of positive (negative) externalities.

It is easy to verify that the life cycle patterns of individual consumption and health expenditure follow as

$$\dot{c} = \frac{u_c(c)}{u_{cc}(c)}(\rho - r), \quad (20)$$

$$\dot{h} = -\frac{\mu_{ha} + \mu_{h\bar{H}}\dot{\bar{H}}}{\mu_{hh}} - \frac{\mu_h}{\mu_{hh}} \left( (r + \mu) - \frac{v}{\psi^P} \right). \quad (21)$$

Thus, individual consumption follows the same rule the planner would choose. However, the values for  $c$  are not equal to those within social optimum. The differences are implied by those arising with regard to the life cycle allocation of health expenditure. Comparison with (17) shows that individuals fail to take into account the fact that health expenditure should (i) decrease in line with the loss of the current year's value of the individual's net contribution to the externality  $\frac{\eta^H}{u_c}(h - \bar{H})$  and (ii) change in line

with changes in the value of the externality  $\left(\frac{\eta^H}{u_c}\right)$ .

In order to attain the first best allocation the social planner can introduce the following transfer scheme. Let  $\tau(a, t_0 + a)$  denote

a (net) subsidy on each unit of private health care spending or, equivalently, on each unit of private health care consumed. Hence, for each unit of care the individual only spends an amount of  $1 - \tau(a, t_0 + a)$ . In order to balance the budget in expected terms the government levies a (net) lump-sum tax equal to the amount  $\tau(a, t_0 + a)h^*(a, t_0 + a)$ , where  $h^*(a, t_0 + a)$  corresponds to the socially optimal level of health expenditure for an individual aged  $a$  at time  $t$ . Note that the lump-sum transfer is entirely exogenous to individual decision making.<sup>24</sup> The individual asset dynamics under the transfer scheme can then be written as follows

$$\dot{A}(a) = (r + \mu(a, h(a), \bar{H}(t_0 + a)))A(a) + y(a) - c(a) - h(a) + \tau(a, t_0 + a)(h(a) - h^*(a, t_0 + a)).$$

From now on we consider the effects of optimal transfers at population level. Therefore we switch back to the  $(a, t)$  notation used in Sections 2 and 3. From the FOCs of the social welfare problem (SW) and of the individual choice problem (IC) we then obtain

$$\begin{aligned} 1 &= -\psi^S(a, t)\mu_h(a, h^*(a, t), \bar{H}^*(t)) + \frac{\eta^H(t)}{u_c(c^*(a, t))} \\ &= -[\psi^S(a, t) + \Theta(a, t)]\mu_h(a, h^*(a, t), \bar{H}^*(t)) \end{aligned} \quad (\text{SW})$$

$$1 = -\psi^P(a, t)\mu_h(a, h(a, t), \bar{H}(t)) + \tau(a, t), \quad (\text{IC})$$

respectively. Here, variables superscribed with ‘\*’ correspond to the socially efficient solution. Combining the two first-order conditions yields

$$\begin{aligned} \tau(a, t) &= \psi^P(a, t)\mu_h(a, h(a, t), \bar{H}(t)) - [\psi^S(a, t) + \Theta(a, t)] \\ &\quad \times \mu_h(a, h^*(a, t), \bar{H}^*(t)). \end{aligned} \quad (22)$$

If the transfer induces optimal health expenditure for all ages, then mortality rates and consumption levels correspond to the socially optimal levels as well. It follows that the PVOL component of the SVOL and the PVOL within the individual's problem are equalised within each age-group. The following is then easily verified (see Appendix E).

**Proposition 2.** *The transfer*

$$\tau^*(a, t) = 1 - \frac{\psi^P(a, t)}{\psi^P(a, t) + \Omega(a, t) + \Theta(a, t)} \quad (23)$$

with  $\Theta(a, t)$ ,  $\psi^P(a, t)$  and  $\Omega(a, t)$  as defined in (10), (11) and (13), respectively, induces the optimal allocation  $\{h(a, t) = h^*(a, t), c(a, t) = c^*(a, t)\}$  for all  $(a, t)$ .

In order to provide some intuition about the transfer, we observe<sup>25</sup>

$$\begin{aligned} \tau^*(a, t) &\leq 1 \\ \tau^*(a, t) &\geq 0 \iff \Theta(a, t) + \Omega(a, t) \geq 0. \end{aligned}$$

Thus, the transfer constitutes a subsidy on health care, with  $\tau^*(a, t) \in [0, 1]$ , if and only if  $\Theta(a, t) + \Omega(a, t) > 0$ . Typically, we would expect this to be the case with positive externalities, where  $\Theta(a, t) > 0$ . Indeed, a sufficient condition for a subsidy is then given by  $\Omega(a, t) \geq 0$ , where an individual continues to be a net

<sup>24</sup> Here, we assume that the lump-sum transfer is age-specific. Note, however, that in the presence of a perfect annuity market the individual is indifferent as to when the lump-sum transfers are paid/received. One could thus equally well conceive a solution, where the individual pays a constant annual tax (or premium)  $\bar{\tau} = \frac{\int_0^\omega e^{-\rho a} M(a, t_0 + a) \tau(a, t_0 + a) h^*(a, t_0 + a) da}{\int_0^\omega e^{-\rho a} M(a, t_0 + a) da}$  or even a once-and-for-all tax amounting to

$\tau_0 = \int_0^\omega e^{-\rho a} M(a, t_0 + a) \tau(a, t_0 + a) h^*(a, t_0 + a) da$ .

<sup>25</sup> Note that  $\psi^P(a, t) + \Omega(a, t) + \Theta(a, t) = -\mu_h(a, t)^{-1} > 0$ .

contributor towards the positive externality over its remaining life time. Note, however, that this is not always guaranteed:  $\Omega(a, t) < 0$  cannot be ruled out if over their remaining lifetime individuals spend less on health care than the population average.

In the following, let us focus on the intuitive case, where for positive externalities, we have  $\tau^*(a, t) \in [0, 1]$ . Here, the transfer rate represents the share of the full social value of health care that is not internalised by the individual. The optimal transfer is then an increasing function of the current value of the externality,  $\Theta(a, t)$ , and the value of future contributions towards the externality,  $\Omega(a, t)$ . If externalities do not play a role, we have  $\mu_{\bar{H}}(a, t) \rightarrow 0$  and, as is readily verified from (10) and (13),  $\Theta(a, t) + \Omega(a, t) \rightarrow 0$ . Obviously, private choice then approaches the social optimum and  $\tau^*(a, t) \rightarrow 0$ . In the other polar case mortality can only be reduced through collective expenditure  $\bar{H}(t)$ , i.e. we have  $\mu_h(a, t) \rightarrow 0$ . In this case,  $\Theta(a, t) \rightarrow \infty$  and thus  $\tau^*(a, t) \rightarrow 1$ , implying that individuals receive health care free of charge at the point of use but have to pay a lump-sum tax  $\tau^*(a, t)h^*(a, t) \rightarrow h^*(a, t)$ . Indeed, this is precisely the solution we would expect under circumstances where survival is a ‘pure’ public good. The intermediate cases follow immediately.

If  $\Theta(a, t) + \Omega(a, t) < 0$  the transfer constitutes a tax on health expenditure:  $\tau^*(a, t) < 0$ . While this is the plausible outcome in case of negative externalities, where  $\Theta(a, t) < 0$ , again this is not always guaranteed. Individuals for whom health expenditure lies below the average over their remaining life time are no longer contributors towards the externality (but rather recipients) and therefore receive a tax relief, where  $\Omega(a, t) > 0$ . Again, focusing on the intuitive case, where  $\tau^*(a, t) < 0$  for  $\Theta(a, t) < 0$  it is worth noting that the resulting tax is, nevertheless, no mirror image of the subsidy paid in the presence of positive externalities. In particular, we note that  $\tau^*(a, t) > -1$  if and only if  $\psi^P(a, t) > -2[\Theta(a, t) + \Omega(a, t)]$ . Hence, if the negative externality becomes strong enough, tax rates in excess of  $-1$  cannot be ruled out. In the limiting case, where  $\psi^P(a, t) + \Omega(a, t) + \Theta(a, t) \rightarrow 0$  taxation at (close to) infinite rates effectively excludes the individual from the consumption of health care.

Note that we can also interpret the transfer scheme in the context of health insurance: In the presence of positive externalities individuals pay an annual premium amounting to  $\tau^*(a, t)h^*(a, t)$  and care is then provided subject to a co-payment equal to  $1 - \tau^*(a, t)$ . In reality, of course the design of health insurance contracts follows the trade-off between equalising consumption across different health states and containing ex-post moral hazard, i.e. the excess consumption of health care free of charge (see e.g. Zweifel et al., 2009). While a model with uncertain health states (as measured e.g. by stochastic period-by-period mortality  $\mu(a, \cdot)$ ) is beyond the scope of this paper, it can nevertheless be understood intuitively how health insurance should be designed in the joint presence of uncertainty and externalities. With positive externalities the policy-maker seeks to stimulate demand for health care even under certainty and, therefore, generates some ‘base-line moral hazard’ by setting  $\tau^*(a, t) > 0$ . In the presence of uncertainty additional moral hazard is then traded-off against insurance, implying that the transfer be set in excess of  $\tau^*(a, t)$ . From the perspective of externalities, this implies an over-internalisation. In contrast, in the presence of negative externalities the policy-maker seeks to discourage the consumption of health care even beyond the level that would be optimal without insurance, implying  $\tau^*(a, t) < 0$  and, thus, a co-payment  $1 - \tau^*(a, t) > 1$ . In the presence of uncertainty, the insurance motive would require the co-payment to be reduced, thus implying under-internalisation. Note, however, that it is not clear a priori whether insurance with a co-payment below 1 should be provided. Indeed, negative externalities militate against the introduction of health insurance.

Finally, we consider the age-time profile of the transfer rate. Noting that in the social optimum (9) we have  $\Theta(a, t) = \zeta(a, t)\psi^S(a, t)$  with

$$\zeta(a, t) := \frac{\eta^H}{u_c(c)} \left( 1 - \frac{\eta^H}{u_c(c)} \right)^{-1} \quad (24)$$

we can express the optimal transfer rate as

$$\begin{aligned} \tau^*(a, t) &= 1 - \frac{\psi^P(a, t)}{[1 + \zeta(a, t)]\psi^S(a, t)} \\ &= 1 - \frac{\psi^P(a, t)}{[1 + \zeta(a, t)][\psi^P(a, t) + \Omega(a, t)]}. \end{aligned} \quad (25)$$

Hence, the optimal transfer decreases in the weighted ratio of PVOL to SVOL, where the weight  $[1 + \zeta(a, t)]^{-1}$  measures the direction and magnitude of the externality. The following can then be verified (see Appendix F).

$$\begin{aligned} \tau_a^* + \tau_t^* &= -(1 - \tau^*) \frac{\Omega}{\psi^P + \Omega} \left[ \frac{\psi_a^P + \psi_t^P}{\psi^P} - \frac{\Omega_a + \Omega_t}{\Omega} \right] \\ &\quad - \frac{\psi^P}{\psi^S} \frac{\zeta}{1 + \zeta} \frac{u_{cc}(c)}{u_c(c)} (c_a + c_t) + \frac{\psi^P}{\psi^S} \frac{\eta_t^H}{u_c(c)}. \end{aligned} \quad (26)$$

The change of the optimal transfer rate with age and time is subject to two major drivers. The first term on the RHS captures the impact on the transfer rate of the change in PVOL as opposed to SVOL for a given weight  $[1 + \zeta(a, t)]^{-1}$ . Naturally, an increase of PVOL relative to SVOL should reflect a decline in the transfer rate. Recalling that SVOL contains PVOL,  $\psi^S = \psi^P + \Omega$ , it follows that the change of PVOL is offset against the change in the net value of future contributions to the externality,  $\Omega$ .<sup>26</sup> Consider, for instance, a positive externality and assume that  $\Omega > 0$ , reflecting a positive net contribution in the future. Noting that  $-(1 - \tau^*) < 0$  for all  $\tau^*$ , it follows that the transfer rate is then bound to decrease as long as the change in PVOL exceeds the change in the net value of future contributions. Typically, the PVOL declines with age from a certain age onwards (see e.g. Shepard and Zeckhauser, 1984; Murphy and Topel, 2006). This should trigger an increase in  $\tau^*$  as long as the value of future contributions increases (or decreases by less), implying a growing shortfall of individual health expenditure from the optimal level. In contrast, if the value of future contributions falls faster than the PVOL, then the transfer rate should be adjusted downwards. The numerical analysis in Section 5 provides further illustration of the age-profiles of the various value measures and the transfer rate for the case of both positive and negative externalities.

The second and third terms on the RHS of (26) reflect changes in the weight  $[1 + \zeta(a, t)]^{-1}$  as driven by the underlying changes in the value of the externality  $\frac{\eta^H}{u_c(c)}$ . Consider a positive externality, where  $\zeta(a, t) > 0$ . By increasing the value of the externality an increase (decrease) in consumption with age,  $c_a + c_t > (<)0$ , would then trigger an increase (decrease) in the transfer rate.<sup>27</sup> Finally, the transfer increases if the shadow price of the externality increases over time. This would be the case, for instance, if the impact of average health expenditure on mortality,  $\mu_{\bar{H}} < 0$ , was to increase in magnitude.

<sup>26</sup> In Appendix F we verify that  $\frac{\psi_a^P + \psi_t^P}{\psi^P} - \frac{\psi_a^S + \psi_t^S}{\psi^S} = \frac{\Omega}{\psi^P + \Omega} \left[ \frac{\psi_a^P + \psi_t^P}{\psi^P} - \frac{\Omega_a + \Omega_t}{\Omega} \right]$ .

<sup>27</sup> Naturally, by (20) we have  $c_a + c_t \geq 0$  if and only if  $\rho \leq r$ . Thus, ultimately the impact of consumption on the transfer rate is driven by the interest rate relative to the rate of time preference.



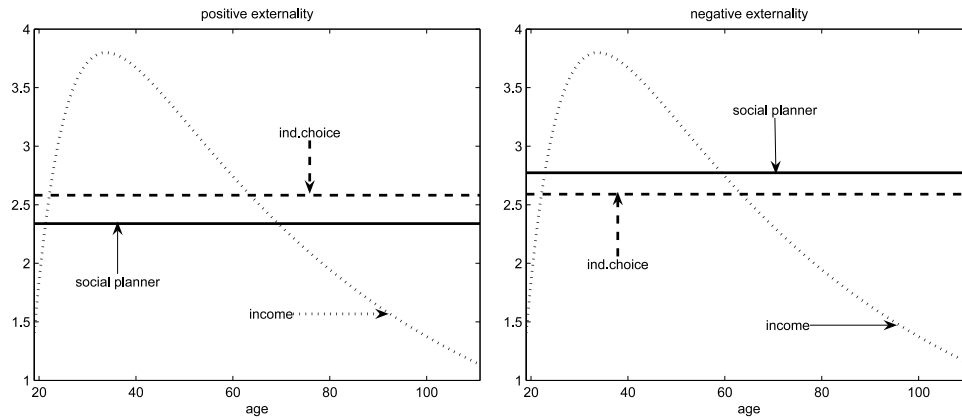


Fig. 2. Consumption in the social planner and the individual choice model (in ten thousands of dollars).

## 5. Numerical analysis

This section illustrates numerically the results derived in the previous sections. We model mortality according to the proportional hazard model (see Kalbfleisch and Prentice, 1980)

$$\mu(a, h(a, t), \bar{H}(t)) = \tilde{\mu}(a)\phi(a, h(a, t), \bar{H}(t)),$$

where  $\tilde{\mu}(a)$  denotes the base mortality rate (effective in the absence of health care) and where  $\phi(a, h(a, t), \bar{H}(t))$  describes the impact of individual and collective health spending. While there is little evidence to guide our choice of the function  $\phi(a, h(a, t), \bar{H}(t))$ , it strikes us as reasonable to assume the following properties:  $\phi_h < 0$ ,  $\phi_{hh} > 0$ ,  $\phi_{ha} > 0$ ;  $\phi(a, 0, 0) = 1$  ( $\forall a$ ) and  $\phi_h(a, 0, 0) = -\infty$  ( $\forall a$ ). Specifically, we let

$$\phi(a, h(a, t), \bar{H}(t)) = 1 - \sqrt{\frac{h(a, t) - \omega}{z} \frac{a - \omega}{1 - \omega}} \pm \sqrt{\frac{\bar{H}(t)}{z'}}. \quad (27)$$

with  $z \geq 0$  and  $z' \geq 0$ . The effectiveness of individual health care is decreasing over age up to the point where it becomes entirely ineffective for  $a = \omega$ . The effect of average expenditure is unrelated to age. Furthermore, for this specification we assume that the marginal effects of individual and aggregate expenditure are unrelated. Finally, note that positive and negative externalities are assumed to be of equal strength as measured by  $|\phi_{\bar{H}}| = (\sqrt{z'\bar{H}(t)})^{-1}$  for any given level of aggregate expenditure  $\bar{H}$ .

Per period utility is specified as

$$u(c(a, t)) = b + \frac{c(a, t)^{1-\sigma}}{1-\sigma}.$$

Our numerical exercise is based on the following data. To generate the age-dependent income schedule,  $y(a)$ , we take from Hall and Jones (2007) the average US-income for 1995 and apply to it a generic age-specific productivity schedule from Skirbekk (2005). The latter is based on productivity estimates constructed as a weighted average over 6 age-dependent abilities (numerical ability, managerial ability, clerical perception, finger dexterity, manual dexterity, experience). Consequently the productivity profile does not represent the productivity for a particular profession, but the average over (more or less) all of them. The maximum life span  $\omega$  is set equal to 110. US mortality data have been taken from the Human Mortality Database (0000) for the years 1990–2000. All our results are calculated for a steady state with a stable population.

We then select parameter values that correspond loosely to the numerical analysis in Hall and Jones (2007). Specifically, we set  $z = 10$  and  $z' = 30$  in the health production function (27) and  $b = 6$  and  $\sigma = 1.5$  in the period utility function. Further,

we set  $\rho = r = 0.06$ . While the interest rate reflects a real return to savings that is also assumed in Hall and Jones (2007), setting the rate of time preference at the same level helps us to characterise the effects of the externalities on the age-profiles of the variables of interest. Specifically, for a constant consumption level over the life cycle the various age-dependent expressions relating to the VOL are not blurred by changes in the marginal utility of consumption over the life course. For this specification we obtain levels of consumption, average health expenditure, a health share of about 15.7, and a statistical VOL for the age-group 35–39 of \$3 million that are broadly comparable to the values in Hall and Jones (2007).<sup>28</sup>

Fig. 2 plots consumption against the income profile (the bell shaped curve). For  $\rho = r$  it is optimal both for the individual and the social planner to smooth consumption over all ages. Thus, as is typical for life cycle models with hump-shaped productivity, individual consumption tends to fall short of income in the youngest and oldest ages and exceeds it over a range of intermediate ages. A comparison between individual and socially optimal consumption reveals that, as expected, individual consumption is too high (low) in the case of positive (negative) externalities. Correspondingly, too little (too much) is spent on health care in the case of positive (negative) spillovers, as is graphed in Fig. 3. More generally, and for any of the cases, the age-profile of health expenditure is hump-shaped. Due to very low base mortality, for the individual there is little to be gained from health care at the youngest ages. While with advancing age the increase in base mortality renders the purchase of health care more and more effective, this is eventually offset for the highest ages, where a falling PVOL and age-related declines in effectiveness lead to a drop in expenditure.<sup>29</sup> This notwithstanding, from a social point of

<sup>28</sup> Hall and Jones (2007): table (II) choose a far higher value of flow utility,  $b = 30.53$  (their scenario 3). Their selection of  $b$  is aimed at obtaining a VOL for the age-group 35–39 of \$3 million for a model in which income is age-independent. In our model income is hump-shaped in age and peaks within the age-band 35–39. Therefore, the VOL tends to be higher due to the human wealth associated with the remaining flow of above average income. With the VOL also increasing in flow utility, we can reduce the latter to a lower value  $b = 6$ .

<sup>29</sup> As a referee has pointed out, a hump-shaped profile of health expenditure stands in contrast to the observation that in most countries (including the US) health expenditure strictly increases with age. The difference arises as our expenditure patterns follow the statistical or *ex-ante* VOL, which typically decreases from some age onwards (see e.g. the numerical exercises in Shepard and Zeckhauser, 1984; Murphy and Topel, 2006). As Philipson et al. (2010) argue, however, real health expenditure is driven by the *ex-post* VOL once a life threatening condition has materialised. At this point individuals are typically willing to spend a manifold of the *ex-ante* VOL. The bunching of life threatening situations at high ages then implies the increasing spending pattern. As our analysis is predominantly of a normative nature and not targeted at a realistic VOL per se but rather how it should be amended in the presence of externalities, the discrepancy between our simulated expenditure and real-world expenditure is of minor consequence.

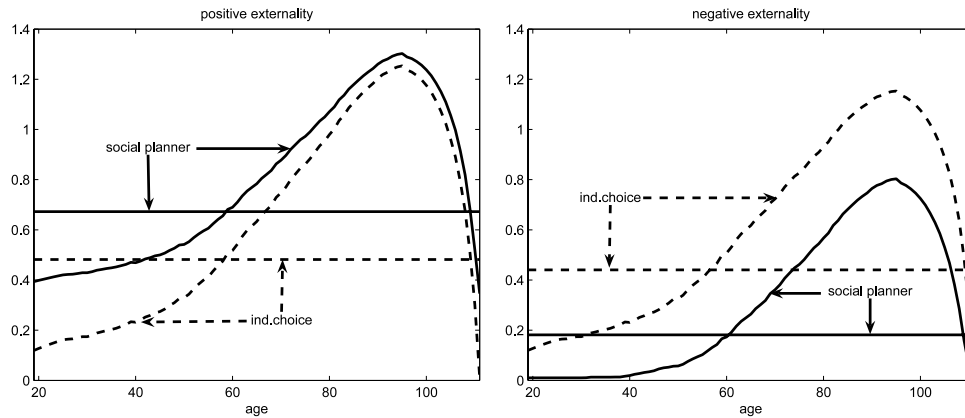


Fig. 3. Age-specific and average (flat line) expenditures in the social planner and the individual choice model (in ten thousands of dollars).

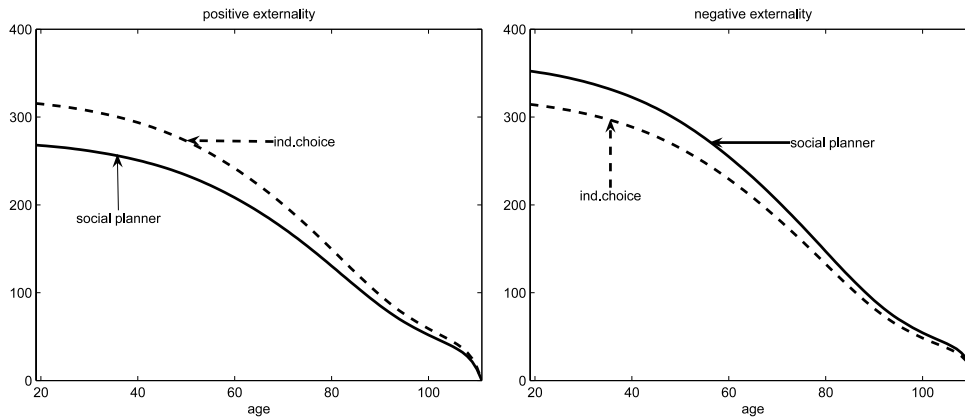


Fig. 4.  $\psi^P$  and  $\psi^S + \Theta$  (in ten thousands of dollars).

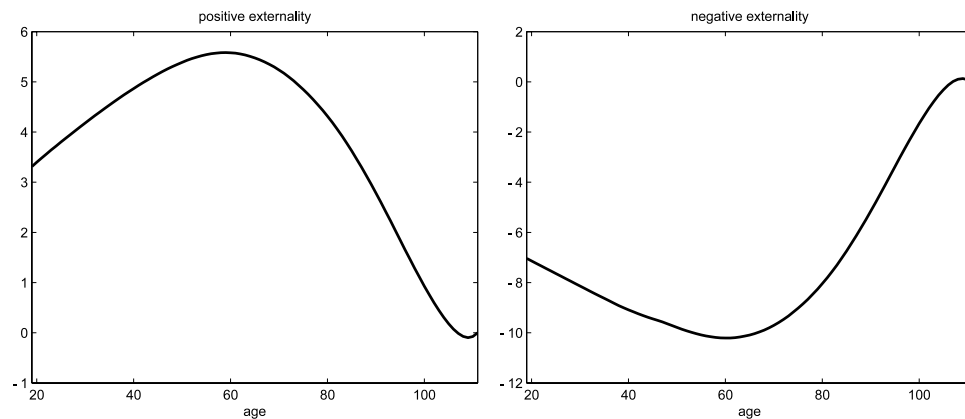


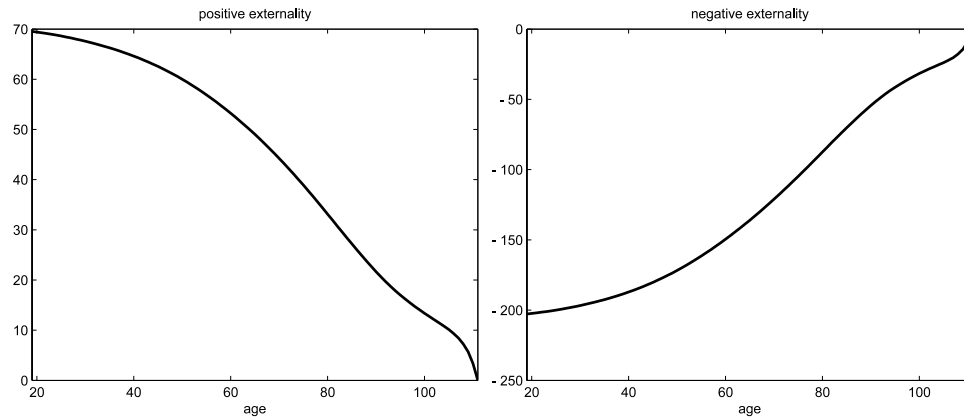
Fig. 5.  $\Omega$ -term (in ten thousands of dollars).

view it is optimal in the presence of positive spillovers to maintain significant spending levels even for the youngest ages in order to ensure sufficient contributions towards aggregate expenditure.

Fig. 4 compares the total private incentive, i.e. the PVOL  $\psi^P$ , in the absence of taxation vs the full social incentive,  $\psi^S + \Theta$ , which we will also refer to as full social value of health care. For our specification both the PVOL and the full social value of health care fall with advancing age. The under- (over-) spending on health in the presence of positive (negative) spillovers is mirrored in the valuations, where PVOL exceeds (falls short of) the social value of health care in the presence of positive (negative) spillovers. Noting from Figs. 5 and 6 (left panel) that in the case of positive externalities the non-private part of the social valuation  $\Omega + \Theta$  is positive for almost every age, it follows that in the

presence of positive externalities the PVOL component of the SVOL (given optimal health spending  $h^*$ ) must fall short of the PVOL in the presence of suboptimal spending. This reflects the fact that individual mortality is 'excessive' when positive spillovers are not internalised, leading to a substantive premium in the willingness to pay for survival. The converse is true in the presence of negative spillovers.

Fig. 7 plots the net increase of life expectancy at age  $a$  that is attainable if the optimal pattern of health expenditure is induced. In the presence of positive externalities (the upper graph), for instance, the life expectancy at age 20 would increase by about 8 months if individuals could be induced to spend optimally. This stands in contrast to the case of negative externalities, where socially optimal behaviour leads to a reduction in life expectancy

Fig. 6.  $\Theta$ -term (in ten thousands of dollars).

up to age 60 and to a modest increase only afterwards. This is the case because from a social point of view, it is optimal to spend close to nothing up to age 40. Note, however, that in contrast to the case with positive externalities the impact on life expectancy of an internalisation of negative externalities is small. The reason is that individual efforts to lower mortality and the externality are to a large extent neutralising each other. Socially desirable reductions in health care do not trigger a large increase in mortality as the negative impact of the externality is diminished in parallel. In the presence of positive externalities, in contrast, additional individual efforts combine with positive spillovers to trigger a sizeable increase in life expectancy.<sup>30</sup>

The comparatively large impact of differential health spending on mortality in the presence of positive as opposed to negative externalities is also evident in the resulting consumption patterns. Consulting Fig. 3 we see that the extent of over-spending on health in the case of negative externalities is more pronounced than the extent of under-spending in the case of positive externalities. This should translate into a gap between socially and individually optimal consumption that is larger in the case of negative externalities. Inspection of 2, however, shows that the gap in consumption is in fact smaller. Indeed, in the presence of positive externalities, socially optimal consumption is lower both due to larger expenditures on health and due to the fact that consumption needs to be spread across additional life years with low income (thus forcing lower levels of per-period consumption).<sup>31</sup> In the presence of negative spillovers, the effect through changes in life expectancy is less pronounced, so that the gap in consumption is more reflective of excessive individual spending on health care. Indeed, in the presence of negative spillovers, the individual's efforts to curb own mortality are ineffective to large extent and only lead to lower consumption: a treadmill-effect.

Fig. 5 develops the age-profile of the net value of an individual's contributions towards the externality over its remaining life time. This value is positive (negative) for all but the highest ages in the case of positive (negative) spillovers. Surprisingly, perhaps, this implies that although individual spending levels exceed the population average only for a range of intermediate life years, in terms of discounted value, the individual tends

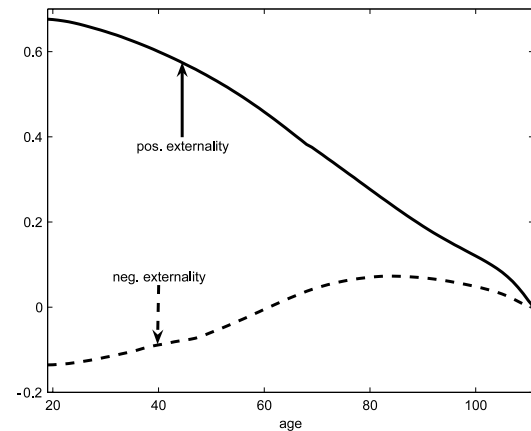


Fig. 7. Difference in life-expectancy.

to be a net contributor towards the externality for all but the highest life years. The hump-shaped spending pattern implies a maximum (minimum) value of future contributions in the case of positive (negative) externalities. As individuals spend less than the population average within their last few life years, the value of future contributions turns negative (positive) in the case of positive (negative) externalities and eventually approaches zero. However, owing to only a few life years left at this point, the absolute values are very low. Finally, we note that in comparison to the PVOL and the current value of the externality,  $\Theta$ , the value of future contributions is rather modest for all ages. Thus, what is predominantly driving the wedge between PVOL and the social value of health expenditure is the current value of the externality.

The latter is plotted in Fig. 6 and has a more straightforward shape. Recall that for an optimal expenditure path  $h^*(a)$  it must be true that  $\Theta(a) = \zeta \psi^S(a)$ , where for a steady-state with  $\rho = r$ , the weight  $\zeta$ , as defined in (24), is a positive (negative) constant in the presence of positive (negative) externalities. Then, as  $\psi^S(a)$  is declining with age, so is the absolute value of  $\Theta(a)$ . What is striking, however, is that in absolute terms, the current value of a negative externality is almost three times that of a positive externality. This is predominantly reflecting higher overall levels of mortality (even under optimal investment) in the presence of negative externalities, which in turn lead to a higher SVOL. Recall from (10) that  $\Theta(a)$  includes the aggregated SVOL of all age groups affected by the externality. Thus, the higher mortality under negative spillovers translates systematically into a higher current valuation of the externality.

Finally, consider the age-profile of the optimal transfer, as in Fig. 8. A first inspection shows that, unsurprisingly, the transfer constitutes a subsidy (tax) in the presence of positive (negative)

<sup>30</sup> Note that the asymmetry in the effects of positive as opposed to negative externalities is, indeed, systematic and not down to our numerical specification (which would support a symmetric outcome). Symmetry of the effects across the two type of externalities would require that in the case of negative externalities both average and individual health expenditure raise mortality. However, this would not make sense from a modelling point of view.

<sup>31</sup> Naturally, this effect would reverse if income was to exceed expenditure for the life years gained.

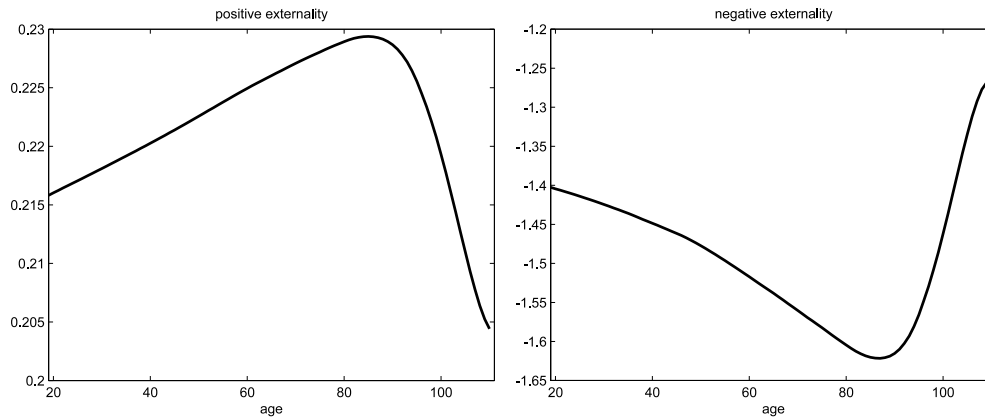


Fig. 8. Transfer.

spillovers. Noting that  $\zeta$  (and, thus,  $\frac{\eta^H}{u_c}$ ) is a constant, it follows from (26) that the age-profile of the transfer is entirely driven by changes of the PVOL and the discounted value of future contributions towards the externality,  $\Omega$ . In the case of positive externalities, the decline in the PVOL suggests a tendency towards an increase in the transfer. Indeed, up to around age 60 this trend is complemented by the increase in the value of future contributions, both widening the wedge between private and social incentives. For higher ages, however, a sharp decline in the value of future contributions becomes dominant, thus leading to a closing gap in incentives and, thus, to a reduction in the subsidy. Surprisingly perhaps, in the case of negative externalities, the decline in the PVOL actually triggers an increase in the tax. Recall from the first equality in (25) that the tax should decrease if and only if the PVOL falls at a rate in excess of the SVOL. From the second equality it can now be checked that for the case of negative externalities, where  $\Omega < 0$ , a reduction of the PVOL by some amount  $d\psi^P < 0$  implies that  $\frac{d\psi^S}{\psi^S} = \frac{d\psi^P}{\psi^S} < \frac{d\psi^P}{\psi^P} < 0$ . Thus, the SVOL declines at a higher rate than the PVOL, implying an increasing tendency towards excess spending. The tax should therefore be increased with age, an effect that is reversed only for ages above the mid 80s, where a strong decline in the (absolute) value of the future contributions towards the externality now generates a tendency towards lower taxation.

A further comparison between the transfer under positive and negative externalities reveals that for negative externalities the absolute value of the transfer tends to be greater by a little less than an order of magnitude. This difference reflects two distinguishing features of the two types of externalities. First, in the presence of positive externalities the PVOL is complemented by a term  $\Theta + \Omega > 0$  to form the full social value of health care. But then, the PVOL is bound to be smaller than the full social value of health care (justifying at best subsidisation at rate 1). In contrast, for negative externalities the social incentive embraces a term  $\Theta + \Omega < 0$  offsetting the PVOL, implying that the full social value of care is necessarily smaller than the PVOL. While in and of itself this does not imply high rates of taxation, it shows that, to some extent, the policy-maker needs strong taxes to reverse private incentives, whereas in the case of positive externalities, even moderate subsidies complement the private incentive. Second, as was discussed previously, the current value of the externality (in absolute terms)  $|\Theta|$  is significantly higher in the presence of negative externalities. This reflects the higher levels of mortality in the presence of negative externalities, translating into a larger gap between private and social incentives and, therefore, calling for a stronger policy.

## 6. Conclusions

We provide a framework for assessing the efficiency of the individual consumption of health care within a continuous-time life cycle framework assuming that health care curbs own mortality but also induces a spill over effect on other person's survival. To this end we combine an age-structured optimal control model at the population level with a conventional life cycle model.

We show analytically how the planner (but not the individual) incorporates in her decision making the value of the externality. Referring to the concept of the willingness to pay for a small reduction of the mortality rate, i.e. the value of life, we compare the social and private incentive to spend on health care. We show that the social value of life comprises the private value of life and, in addition, the individual's (net) future contribution towards social welfare through the spillovers from individual consumption of health care. Adding to the social value of life the current value of the externality, as generated by current individual health care spending, results in the full social value of health care, which in optimum is equalised with the effective cost of life saving. We derive analytically the age-profiles of consumption (following the usual pattern) and, more interestingly, of health expenditure. With regard to the latter we can trace out explicitly how it develops depending on changes related to the effects of age and aggregate expenditure on the 'effectiveness of care' and to changes in the underlying valuations.

From a comparison of the individual and social incentive we derive a transfer scheme that induces the individual to consume health care at the socially efficient level. The optimal transfer rate amounts to the proportion of the external value of health care relative to its full social value (including in addition the private value of life). We examine analytically the properties of this transfer, in particular its age dependency, both for the case of positive and negative externalities. The age-pattern of the transfer depends, in particular, on the change of the private value of life relative to the value of future contributions to the externality, as moderated by the direction of the externality.

We illustrate the main findings of our model through numerical illustrations based on US data. The nature of the externality turns out to have distinct consequences for the pattern of mortality. In the presence of positive spillovers mortality can be reduced sizably with corresponding increases in life expectancy. However, owing to their under-spending individuals fail to realise a significant share of these gains in life expectancy. In contrast, when spillovers are negative, no substantial reductions in mortality below the baseline can be attained. Here, the inefficiency of individual behaviour is manifested in a tread-mill effect, where individuals over-spend on health care without great effect and, thereby, forego consumption. Finally, we plot age-profiles of the optimal transfer



rate, of the private value of life, the full social value of health care, as well as of the current and future values of individual health spending.

Our results indicate that for negative externalities the value of the transfer tends to be greater as compared to the transfer under positive externalities reflecting higher overall mortality and a difficulty to reverse rather than just complement private incentives. Furthermore, both the subsidy and the tax increase in absolute value up to an age 80+, only to decline for the highest age groups. Thus, the transfer peaks at around the age (or shortly before) with highest medical spending. This age-pattern may have rather strong policy implications. Suppose the transfer is implemented as a co-payment to health insurance. In the presence of positive externalities the co-payment is then reduced with age up to the ages with highest spending, a pattern which should make this an attractive policy option from the individual's point of view. In contrast, negative externalities would require an increase of co-payments with age with a peak at the ages with highest spending: clearly a very unappealing option. These arguments should be understood to be illustrative, not the least because they rely on the stylised assumption that there is an exclusive presence of either positive or negative externalities.

More generally, it has been our main objective to provide a modelling framework to analyse the efficiency of individual life cycle behaviour, to present the critical elements of such an analysis and to illustrate the channels of transmission by which direct period effects and effects through changes in the life expectancy impact on life cycle choice. In order to facilitate the representation as much as possible we have therefore adopted a number of simplifying assumptions regarding the nature of the externalities. In particular, by assuming that the spillovers flow through current aggregate health care expenditure we presume (i) that all age-groups contribute in a symmetric way and, more importantly, (ii) that there are no cumulative effects of spending, as would be present in the context of stock externalities. For our numerical analysis, we impose additional assumptions, namely that the marginal productivity of individual health expenditure is unaffected by the externality and that the impact of the externality on mortality is independent of age. Clearly, these assumptions are unrealistic and compromise a direct application of our results to the various examples of real-world externalities that were discussed in the introduction. Nevertheless, we have identified a number of important drivers behind the life cycle transmission of survival-related externalities that generalise to more realistic settings. While there is apparent scope for drawing up more realistic models of life cycle externalities, we leave this to future research.

Finally, our model lends itself to the analysis of other imperfections in individual behaviour. The cross-cohort spillovers that give rise to inefficiency are clearly not only present in the health care sector but also – and perhaps more prominently – in the production and/or consumption of goods. Externalities in production arise with regard to saving towards the accumulation of a (common) capital stock that affects the productivity of (everyone's) labour; and with regard to health or educational investments that increase individual productivity but also the productivity of co-workers. Externalities with regard to consumption arise for many modes of unhealthy consumption (see Forster, 2001, for a life cycle-model of unhealthy consumption without spillovers). Most prominently this relates to smoking which not only raises individual mortality but also the mortality of others. Similar arguments apply, however, to other consumption goods, such as cars, that directly or indirectly lead to the emission of pollutants. As should have become evident from our analysis, such externalities will lead to distortions both through period effects and through effects in overall life expectancy. We would thus envisage a number of interesting applications.

## Appendix A. Proof of Proposition 1

Consider the Hamiltonian in (4). The dynamics of the adjoint variables satisfy:

$$\dot{\xi}_a^N + \xi_t^N = (\rho + \mu(a, h, \bar{H}))\xi^N - u(c) - \xi^A(y - c - h) - \eta^H h - \eta^{\bar{N}} \quad (28)$$

$$\dot{\xi}_a^A + \xi_t^A = (\rho - r)\xi^A \quad (29)$$

$$\eta^H(t) = - \int_0^\omega \xi^N \mu_{\bar{H}}(a, h, \bar{H}) \frac{N}{\bar{N}} da \quad (30)$$

$$\eta^{\bar{N}}(t) = \int_0^\omega \xi^N \mu_{\bar{H}}(a, h, \bar{H}) \frac{HN}{\bar{N}^2} da = -\eta^H(t)\bar{H}(t) \quad (31)$$

together with

$$\xi^N(\omega, t) = 0.$$

From now on, we assume  $T < +\infty$ , which further implies  $\xi^N(a, T) = 0$ . In order to obtain transversality conditions for  $\xi^A$ , we have to consider the conditions  $A(a, 0) = A_0(a)$ ,  $A(a, T) = A_T(a)$ ,  $A(0, t) = 0$  and  $A(\omega, t) = 0$ . For age-specific optimal control models with initial and end state conditions there are no transversality conditions. Thus we ignore  $A(\omega, t) = 0$  and  $A(a, T) = A_T(a)$  and add the terms  $-\lambda \int_0^T e^{-rt} A(\omega, t)^2 dt$  and  $-\lambda \int_0^\omega e^{-rt} (A(a, T) - A_T(a))^2 da$  to the objective function. Thus we obtain  $\xi^A(\omega, t) = -2\lambda A(\omega, t) > 0$  and  $\xi^A(a, T) = -2\lambda (A(a, T) - A_T(a))$  as transversality conditions (implying  $\xi^A(a, t) > 0$  for  $\forall (a, t)$ ). The necessary first order conditions with respect to the controls  $c$  and  $h$  are given by

$$\mathcal{H}_c = u_c(c)N - \xi^A N = 0$$

$$\mathcal{H}_h = -\xi^N \mu_h(a, h, \bar{H})N - \xi^A N + \eta^H N = 0$$

as reported in the text. Combining them, we obtain  $u_c(c) = -\xi^N \mu_h(a, h, \bar{H}) + \eta^H$  or

$$\begin{aligned} \frac{-1}{\mu_h(a, h, \bar{H})} &= \frac{\xi^N(a, t)}{u_c(c)} - \frac{\eta^H(t)}{u_c(c(a, t))\mu_h(a, h, \bar{H})} \\ &= \psi^S(a, t) + \Theta(a, t), \end{aligned} \quad (32)$$

where the second equation follows using the definition of (8) together with

$$\begin{aligned} \Theta(a, t) &:= - \frac{\eta^H(t)}{u_c(c(a, t))\mu_h(a, h, \bar{H})} \\ &= \int_0^\omega \frac{u_c(c(\hat{a}, t))}{u_c(c(a, t))} \psi^S(\hat{a}, t) \frac{N(\hat{a}, t)}{\bar{N}(t)} \frac{\mu_{\bar{H}}(\hat{a}, h, \bar{H})}{\mu_h(a, h, \bar{H})} d\hat{a}. \end{aligned}$$

Here, the second equality follows when inserting from (30) and using  $\xi^N(\hat{a}, t) = u_c(c(\hat{a}, t))\psi^S(\hat{a}, t)$ .<sup>32</sup>

## Appendix B. Proof of Lemma 1

Solving the adjoint equation (28) with the method of characteristics we obtain,

$$\begin{aligned} \xi^N(a, t) &= \int_a^\omega e^{-\int_a^s [\rho + \mu(s', h, \bar{H})] ds'} \\ &\times \left[ u(c(s, t - a + s)) + u_c(c(s, t - a + s)) \right. \\ &\times (y(t - a + s) - c(t - a + s) - h(t - a + s)) \Big] ds \\ &+ \int_a^\omega e^{-\int_a^s [\rho + \mu(s', h, \bar{H})] ds'} \left( \int_0^\omega \xi^N \mu_{\bar{H}}(\hat{a}, h, \bar{H}) \frac{N}{\bar{N}} d\hat{a} \right) \\ &\times (\bar{H}(t - a + s) - h(s, t - a + s)) ds \end{aligned}$$

<sup>32</sup> Note the change in notation from  $a$  to  $\hat{a}$ , where  $a$  refers to the age-group for which the optimality condition holds and  $\hat{a}$  refers to any of the age-groups.

where we make use of the fact that  $\xi^N(\omega, t) = 0$  for every  $t$ . Using the definition in (12) and inserting from (30) and (31) we obtain

$$\begin{aligned}\xi^N(a, t) &= \int_a^\omega e^{-\int_a^s [\rho + \mu(s', h, \bar{H})] ds'} \\ &\quad \times u_c(c(s, t - a + s)) v(s, t - a + s) ds \\ &\quad - \int_a^\omega e^{-\int_a^s [\rho + \mu(s', h, \bar{H})] ds'} \eta^H(t - a + s) \\ &\quad \times [\bar{H}(t - a + s) - h(s, t - a + s)] ds.\end{aligned}$$

Employing the definition in (8) we can then write

$$\begin{aligned}\psi^S(a, t) &= \frac{\xi^N(a, t)}{u_c(c(a, t))} = \int_a^\omega e^{-\int_a^s [\rho + \mu(s', h, \bar{H})] ds'} \\ &\quad \times \frac{u_c(c(s, t - a + s))}{u_c(c(a, t))} v(s, t - a + s) ds \\ &\quad - \int_a^\omega e^{-\int_a^s [\rho + \mu(s', h, \bar{H})] ds'} \\ &\quad \times \frac{u_c(c(s, t - a + s))}{u_c(c(a, t))} \frac{\eta^H(t - a + s)}{u_c(c(s, t - a + s))} \\ &\quad \times (\bar{H}(t - a + s) - h(s, t - a + s)) ds.\end{aligned}$$

Noting that the first order condition for  $c(a, t)$  as in (5) implies  $\frac{u_c(c(s, t - a + s))}{u_c(c(a, t))} = \frac{\xi^A(s, t - a + s)}{\xi^A(a, t)} = e^{(\rho - r)(s - a)}$  and substituting gives

$$\begin{aligned}\psi^S(a, t) &= \int_a^\omega e^{-\int_a^s [r + \mu(s', h, \bar{H})] ds'} v(s, t - a + s) ds \\ &\quad - \int_a^\omega e^{-\int_a^s [r + \mu(s', h, \bar{H})] ds'} \frac{\eta^H(t - a + s)}{u_c(c(s, t - a + s))} \\ &\quad \times (\bar{H}(t - a + s) - h(s, t - a + s)) ds \\ &= \psi^P(a, t) - \int_a^\omega e^{-\int_a^s [r + \mu(s', h, \bar{H})] ds'} \\ &\quad \times \frac{\Theta(s, t - a + s)}{\psi^S(s, t - a + s) + \Theta(s, t - a + s)} \\ &\quad \times (\bar{H}(t - a + s) - h(s, t - a + s)) ds \\ &= \psi^P(a, t) + \Omega(a, t),\end{aligned}$$

where the second equality follows when observing the definition in (11) as well as the relationship  $\frac{\eta^H(t - a + s)}{u_c(c(s, t - a + s))} = \frac{\Theta(s, t - a + s)}{\psi^S(s, t - a + s) + \Theta(s, t - a + s)}$ , following from the second equation in (32).

### Appendix C. Derivation of the dynamics of $h$

From the necessary first order condition expressed in value terms, i.e. (32), we obtain

$$-1 = \mu_h(1 + \zeta)\psi^S \quad (33)$$

where

$$\Theta = \zeta\psi^S \quad (34)$$

$$\zeta := \frac{\eta^H}{u_c} \quad (35)$$

are observed. Total differentiation in direction  $(a, t)$  yields

$$\begin{aligned}h_a + h_t &= -\frac{1}{\mu_{hh}} \left[ \mu_{ha} + \mu_{h\bar{H}} \bar{H}_t \right. \\ &\quad \left. + \mu_h \left( \frac{(\psi_a^S + \psi_t^S)}{\psi^S} + \frac{(\zeta_a + \zeta_t)}{1 + \zeta} \right) \right].\end{aligned} \quad (36)$$

Using  $\psi^S = \psi^P + \Omega$  we obtain  $(\psi_a^S + \psi_t^S) = (\psi_a^P + \psi_t^P) + (\Omega_a + \Omega_t)$ . From (11) and (13) we obtain

$$\begin{aligned}\psi_a^P + \psi_t^P &= (r + \mu)\psi^P - v \\ \Omega_a + \Omega_t &= (r + \mu)\Omega - \frac{\eta^H}{u_c}(h - \bar{H}).\end{aligned} \quad (37)$$

Hence,

$$\frac{\psi_a^S + \psi_t^S}{\psi^S} = (r + \mu) - \frac{1}{\psi^S} \left[ v + \frac{\eta^H}{u_c}(h - \bar{H}) \right]. \quad (38)$$

Furthermore,

$$\begin{aligned}\frac{(\zeta_a + \zeta_t)}{1 + \zeta} &= -\frac{1}{\mu_h \psi^S} \left[ \left( \frac{\eta^H}{u_c} \right)_a + \left( \frac{\eta^H}{u_c} \right)_t \right] \\ &= -\frac{1}{\mu_h \psi^S} \left[ \frac{\eta_t^H}{u_c(\cdot)} - \frac{\eta^H}{u_c(\cdot)} (\rho - r) \right],\end{aligned} \quad (39)$$

where the second equality follows when observing that  $\left( \frac{\eta^H}{u_c} \right)_a + \left( \frac{\eta^H}{u_c} \right)_t = \frac{\eta_t^H}{u_c(\cdot)} - \frac{\eta^H}{[u_c(\cdot)]^2} u_{cc}(c_a + c_t)$ , inserting from (16) and cancelling terms. Inserting in (36) we obtain Eq. (17).

### Appendix D. The individual optimum

The Hamiltonian of the individual problem reads (again omitting  $a$  whenever it is not of particular importance)

$$\begin{aligned}\mathcal{H} &= u(c)M - \lambda_M \mu(a, h, \bar{H})M \\ &\quad + \lambda_A((r + \mu(a, h, \bar{H}))A + y - c - h)\end{aligned}$$

where  $\lambda_M$  and  $\lambda_A$  denote the adjoint variables of the survival probability and individual assets respectively. From the necessary optimality conditions we can derive the following system of adjoint variables:

$$\begin{aligned}\dot{\lambda}_M &= (\rho + \mu(a, h, \bar{H}))\lambda_M - u(c) \\ \dot{\lambda}_A &= (\rho - r - \mu(a, h, \bar{H}))\lambda_A\end{aligned} \quad (40)$$

with the transversality conditions  $\lambda_M(\omega) = 0$  and  $\lambda_A(\omega) = -2\lambda A(\omega)$ , since we implement the terminal condition  $A = 0$  in the same way as in the planner's problem. Thus both adjoint variables are always positive. The necessary first order conditions are

$$\mathcal{H}_c = u_c M - \lambda_A = 0 \quad (41)$$

$$\mathcal{H}_h = -(\lambda_M M - \lambda_A A)\mu_h - \lambda_A = 0.$$

Combining them we obtain

$$\begin{aligned}u_c M &= -(\lambda_M M - \lambda_A A)\mu_h \\ \Leftrightarrow u_c &= -(\lambda_M - u_c A)\mu_h \\ \Leftrightarrow \frac{-1}{\mu_h(a, h, \bar{H})} &= \frac{\lambda_M(a)}{u_c(c(a))} - A(a).\end{aligned}$$

Integrating out (40) and (18) we can rewrite the last condition as

$$\begin{aligned}\frac{-1}{\mu_h(a, h, \bar{H})} &= \frac{1}{u_c(c(a))} \int_a^\omega u(c(s)) e^{-\int_a^s [\rho + \mu(s', h, \bar{H})] ds'} ds \\ &\quad + \int_a^\omega [y(s) - c(s) - h(s)] e^{-\int_a^s [r + \mu(s', h, \bar{H})] ds'} ds \\ &= \int_a^\omega \frac{u_c(c(s))}{u_c(c(a))} \frac{u(c(s))}{u_c(c(s))} e^{-\int_a^s [\rho + \mu(s', h, \bar{H})] ds'} ds \\ &\quad + \int_a^\omega [y(s) - c(s) - h(s)] e^{-\int_a^s [r + \mu(s', h, \bar{H})] ds'} ds \\ &= \int_a^\omega v(s) e^{-\int_a^s [r + \mu(s', h, \bar{H})] ds'} ds = \psi^P(a).\end{aligned}$$

Here, the third equality follows when observing that (41) implies  $\frac{u_c(c(s))}{u_c(c(a))} = \frac{\lambda_A(s)M(a)}{\lambda_A(a)M(s)} = e^{(\rho - r)(s - a)}$  and employing definition (12). The last equality follows by definition (11).

## Appendix E. Proof of Proposition 2

An optimal transfer induces spending levels  $h(a, t) = h^*(a, t)$   $\forall(a, t)$ . In turn, this implies  $\bar{H}(t) = \bar{H}^*(t)$  and, thus,  $\mu_h(a, h(a, t), \bar{H}(t)) = \mu_h(a, h^*(a, t), \bar{H}^*(t))$ . We can therefore, rewrite (22) as

$$\begin{aligned}\tau^*(a, t) &= \psi^P(a, t)\mu_h(a, h^*(a, t), \bar{H}^*(t)) \\ &\quad - [\psi^S(a, t) + \Theta(a, t)]\mu_h(a, h^*(a, t), \bar{H}^*(a, t)) \\ &= [\psi^P(a, t) - \psi^S(a, t) - \Theta(a, t)]\mu_h(a, h^*(a, t), \bar{H}^*(t)).\end{aligned}$$

Furthermore, note that  $h(a, t) = h^*(a, t)$  and  $\mu_h(a, h(a, t), \bar{H}(t)) = \mu_h(a, h^*(a, t), \bar{H}^*(t))$  imply  $c(a, t) = c^*(a, t)$ . Using the definition  $\psi^S(a, t) = \psi^P(a, t) + \Omega(a, t)$  together with  $c(a, t) = c^*(a, t)$  and  $h(a, t) = h^*(a, t)$  we can then write  $\tau^*(a, t) = -[\Omega(a, t) + \Theta(a, t)]\mu_h(a, h^*(a, t), \bar{H}^*(t))$ . Substituting from the first-order condition (15)  $\mu_h(a, h^*(a, t), \bar{H}^*(t)) = [\psi^P(a, t) + \Omega(a, t) + \Theta(a, t)]^{-1}$  and rearranging we then obtain the expression reported in the Proposition.

## Appendix F. Derivation of the dynamics of $\tau$

Taking the total derivative of (25) with respect to  $a$  and  $t$  we obtain

$$\begin{aligned}\tau_a^* + \tau_t^* &= -\frac{\psi^S(\psi_a^P + \psi_t^P) - \psi^P(\psi_a^S + \psi_t^S)}{(1 + \zeta)(\psi^S)^2} + \frac{\psi^P(\zeta_a + \zeta_t)}{(1 + \zeta)^2\psi^S} \\ &= -\frac{\psi^P}{(1 + \zeta)\psi^S} \left[ \frac{(\psi_a^P + \psi_t^P)}{\psi^P} - \frac{(\psi_a^S + \psi_t^S)}{\psi^S} \right] \\ &\quad + \frac{\psi^P(\zeta_a + \zeta_t)}{(1 + \zeta)^2\psi^S}.\end{aligned}\quad (42)$$

The first term on the RHS can be rewritten as follows

$$\begin{aligned}& -\frac{\psi^P}{(1 + \zeta)\psi^S} \left[ \frac{(\psi_a^P + \psi_t^P)}{\psi^P} - \frac{(\psi_a^S + \psi_t^S)}{\psi^S} \right] \\ &= -(1 - \tau^*) \left[ \frac{(\psi_a^P + \psi_t^P)}{\psi^P} - \frac{(\psi_a^P + \psi_t^P) + (\Omega_a + \Omega_t)}{\psi^P + \Omega} \right] \\ &= -(1 - \tau^*) \left[ \frac{(\psi_a^P + \psi_t^P)}{\psi^P} - \frac{\psi^P}{\psi^P + \Omega} \frac{(\psi_a^P + \psi_t^P)}{\psi^P} \right. \\ &\quad \left. - \frac{\Omega}{\psi^P + \Omega} \frac{(\Omega_a + \Omega_t)}{\Omega} \right] \\ &= -(1 - \tau^*) \frac{\Omega}{\psi^P + \Omega} \left[ \frac{(\psi_a^P + \psi_t^P)}{\psi^P} - \frac{(\Omega_a + \Omega_t)}{\Omega} \right].\end{aligned}\quad (43)$$

Using (39) and then (33) and (35) we can rewrite the second term on the RHS in (42) as

$$\begin{aligned}& \frac{\psi^P}{(1 + \zeta)\psi^S} \frac{(\zeta_a + \zeta_t)}{1 + \zeta} \\ &= -\frac{\psi^P}{(1 + \zeta)\psi^S} \frac{1}{\mu_h\psi^S} \left( \left( \frac{\eta^H}{u_c(\cdot)} \right)_a + \left( \frac{\eta^H}{u_c(\cdot)} \right)_t \right) \\ &= -\frac{\psi^P}{\psi^S} \frac{\zeta}{1 + \zeta} \frac{u_{cc}(\cdot)}{u_c(\cdot)} (c_a + c_t) + \frac{\psi^P}{\psi^S} \frac{\eta_t^H}{u_c(\cdot)}.\end{aligned}\quad (44)$$

Combining (43) and (44) we obtain Eq. (26).

## References

- Basu, A., Meltzer, D., 2005. Implications of spillover effects within the family for medical cost-effectiveness analysis. *Journal of Health Economics* 24, 751–773.
- Birchenall, J.A., Soares, R.R., 2009. Altruism, fertility, and the value of children: health policy evaluation and intergenerational welfare. *Journal of Public Economics* 93, 280–295.
- Black, D., Pearson, M., 2002. Average length of stay, delayed discharge, and hospital congestion. *British Medical Journal* 325, 610–611.
- Bolin, K., Jacobson, L., Lindgren, B., 2002. Employer investments in employee health. Implications for the family as health producer. *Journal of Health Economics* 21, 563–583.
- Boucekkine, R., de la Croix, D., Licandro, O., 2004. Modeling vintage structures with PDES: principles and applications. *Mathematical Population Studies* 11, 151–179.
- Clement, J.P., Valdmanis, V.G., Bazzoli, G.J., Zhao, M., Chukmaitov, A., 2008. Is more better? an analysis of hospital outcomes and efficiency with a DEA model of output congestion. *Health Care Management Science* 11, 67–77.
- Cohen, F.L., Tartasky, D., 1997. Microbial resistance to drug therapy: a review. *American Journal of Infection Control* 25, 51–64.
- Cutler, D.M., Miller, G., 2005. The role of public health improvements in health advances: the twentieth-century United States. *Demography* 42, 1–22.
- Davies, J.B., Kuhn, P., 1992. Social security, longevity, and moral hazard. *Journal of Public Economics* 49, 91–106.
- Easterlin, R.A., 1999. How beneficent is the market? a look at the modern history of mortality. *European Review of Economic History* 3, 257–294.
- Ehrlich, I., 2000. Uncertain lifetime, life protection, and the value of life saving. *Journal of Health Economics* 19, 341–367.
- Ehrlich, I., Chuma, H., 1990. A model of the demand for longevity and the value of life extension. *Journal of Political Economy* 98, 761–782.
- Feichtinger, G., Hartl, R.F., Kort, P.M., Veliov, V.M., 2006. Capital accumulation under technological progress and learning: a vintage capital approach. *European Journal of Operational Research* 172, 293–310.
- Feichtinger, G., Tragler, G., Veliov, V.M., 2003. Optimality conditions for age-structured control systems. *Journal of Mathematical Analysis and Applications* 288, 47–68.
- Forster, M., 2001. The meaning of death: some simulations of a model of healthy and unhealthy consumption. *Journal of Health Economics* 20, 613–638.
- Grossman, M., 1972. On the concept of health capital and the demand for health. *Journal of Political Economy* 80, 223–255.
- Hall, R.E., Jones, C.I., 2007. The value of life and the rise in health spending. *Quarterly Journal of Economics* 122, 39–72.
- Human Mortality Database. <http://www.mortality.org/>.
- Johansson, P.-O., 2002. On the definition and age-dependency of the value of a statistical life. *Journal of Risk and Uncertainty* 25, 251–263.
- Kalbfleisch, J.D., Prentice, R.L., 1980. *The Statistical Analysis of Failure Time Data*. Wiley & Sons, New York.
- Keyfitz, B.L., Keyfitz, N., 1997. The McKendrick partial differential equation and its uses in epidemiology and population study. *Mathematical and Computer Modelling* 26, 1–9.
- Kuhn, M., Wrzaczek, S., Oeppen, J., 2010. Recognizing progeny in the value of life. *Economics Letters* 107, 17–21.
- Laxminarayan, R., Brown, G.M., 2001. Economics of antibiotic resistance: a theory of optimal use. *Journal of Environmental Economics and Management* 42, 183–206.
- Murphy, K.M., Topel, R.H., 2006. The value of health and longevity. *Journal of Political Economy* 114, 871–904.
- Murphy, K.M., Topel, R.H., 2007. Social value and the speed of innovation. *American Economic Review* 97, 433–437.
- Phillips, K.A., Luft, H., 1997. The policy implications of using hospital and physician volumes as ‘indicators’ of quality of care in a changing health care environment. *International Journal for Quality in Health Care* 9, 341–348.
- Philipson, T., 2000. Economic epidemiology and infectious diseases. In: Culyer, A.C., Newhouse, J.P. (Eds.), *Handbook of Health Economics*, vol. 1B. Elsevier, Amsterdam.
- Philipson, T., Becker, G.S., 1998. Old-age longevity and mortality-contingent claims. *Journal of Political Economy* 106, 551–573.
- Philipson, T.J., Becker, G., Goldman, D., Murphy, K.M., 2010. Terminal care and the value of life near its end. NBER Working Paper 15649.
- Philipson, T., Mechoulam, S., 2003. Intellectual property & external consumption effects: generalizations from pharmaceutical markets. NBER Working Paper 9598.
- Prskawetz, A., Veliov, V.M., 2007. Age specific dynamic labor demand and human capital investment. *Journal of Economic Dynamics and Control* 31, 3741–3777.
- Rosen, S., 1988. The value of changes in life expectancy. *Journal of Risk and Uncertainty* 11, 285–304.
- Saglam, C., Veliov, V.M., 2008. The role of endogenous vintage specific depreciation on optimal behavior of firms. *International Journal of Economic Theory* 4, 381–410.
- Shepard, D.S., Zeckhauser, R.J., 1984. Survival versus consumption. *Management Science* 30, 423–439.
- Skirbekk, V., 2005. Why not start younger? implications of the timing and duration of schooling for fertility, human capital productivity, and public pensions. IIASA RR-05-002.
- Watson, T., 2006. Public health investments and the infant mortality gap: evidence from federal sanitation interventions on US indian reservations. *Journal of Public Economics* 90, 1537–1560.
- Yaari, M.E., 1965. Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies* 32, 137–150.
- Zweifel, P., Breyer, F., Kifmann, M., 2009. *Health Economics*. Springer, Berlin, Heidelberg.